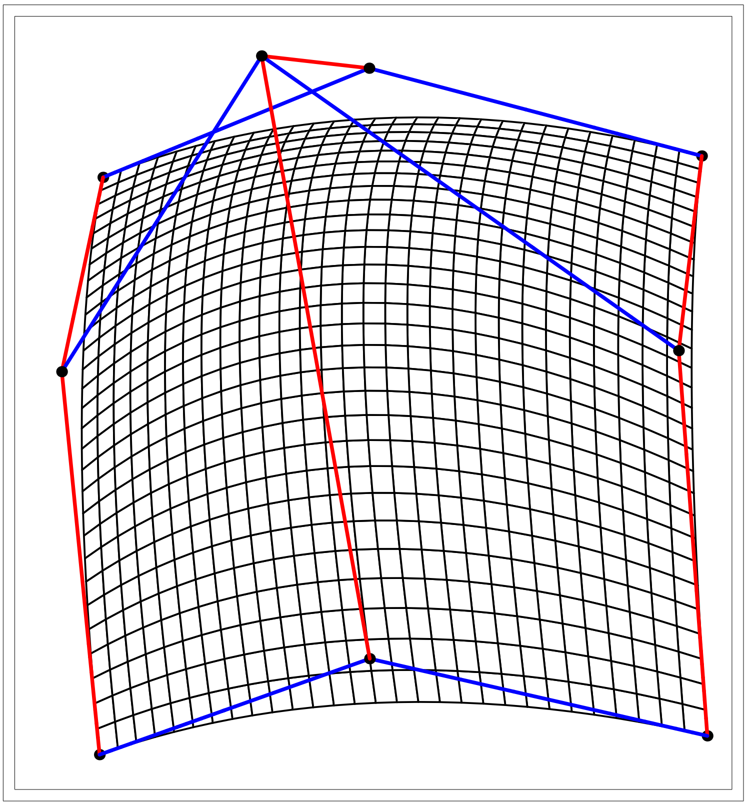
Ángel Montesdeoca Delgado

Geometría Afín y Euclídea



Departamento de Matemática Fundamental
Universidad de La Laguna
2012

Geometría Afín y Euclídea

Ángel Montesdeoca Delgado

(Versión 2.1301181326)

amontes@ull.es http://webpages.ull.es/users/amontes

CONTENIDO

L'algèbre n'est qu'une géométrie écrite; la géométrie n'est qu'une algèbre figurée.
Sophie Germain (1776-1831). Ouvres philosophiques, p. 223
http://www.agnesscott.edu/lriddle/women/germain.htm

TEMA 1
Espacios vectoriales

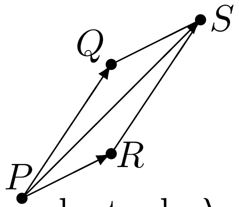
Como el concepto de espacio afín (que se introduce en el §2.1) surge como una abstracción de las operaciones con vectores libres en el espacio ordinario, empezaremos haciendo uso de conocimientos de la Geometría Elemental para introducir el concepto de vectores libres, operaciones con ellos y destacando algunas propiedades de éstas. Ello, nos llevará al concepto algebraico de espacio vectorial y, aunando los hechos considerados con puntos y vectores libres, al de espacio afín.

1.1 Espacios vectoriales. Definición y ejemplos

Hacemos un repaso somero sobre espacios vectoriales: definición, bases, aplicaciones lineales, producto escalar, etc., omitiendo las demostraciones de la mayor parte de los hechos que exponemos, que pueden ser consultadas en cualquier material de Algebra Lineal.

Vectores libres en el espacio ordinario

En el espacio tridimensional ordinario $\mathcal{E}$ se llama *vector fijo* a un par ordenado de puntos (P, Q) y lo representamos por $\overrightarrow{PQ}$; se dice que P es el punto origen del vector y Q es su extremo; si origen y extremo coinciden, se dice que es el vector cero: $\overrightarrow{PP}$. De un vector fijo $\overrightarrow{PQ}$, se dice que tiene la <u>dirección</u> de la recta determinada por P y Q; que su <u>sentido</u> es del de P a Q; y que su <u>módulo</u> es la distancia entre los puntos P y Q. El conjunto de vectores fijos en el espacio ordinario lo representaremos por $\mathcal{F}$.



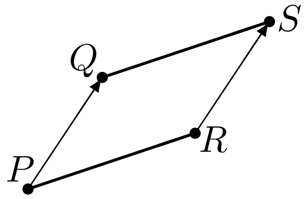
Dados dos vectores $\overrightarrow{PQ}$ y $\overrightarrow{PR}$, con el mismo origen, se llama <u>suma</u> de ellos al vector $\overrightarrow{PS} = \overrightarrow{PQ}+\overrightarrow{PR}$, siendo S el punto tal que $PQSR$ es un paralelogramo.

Dado un vector fijo $\overrightarrow{PQ}$ y un número real λ, se llama <u>producto</u> de λ por $\overrightarrow{PQ}$ al vector $\overrightarrow{PR} = \lambda\overrightarrow{PQ}$, tal que los tres puntos P, Q y R están en una misma recta, con P en el segmento QR o no, si λ es negativo o positivo, respectivamente (si $\lambda = 0$, $R = P$ y se pone $\lambda\overrightarrow{PQ} = \overrightarrow{PP}$). El módulo de $\overrightarrow{PR}$ es producto del módulo $\overrightarrow{PQ}$ por el valor absoluto de λ.

Nótese que la suma de vectores fijos no es una ley de composición en el conjunto $\mathcal{F}$ de todos ellos, pues no se pueden sumar si no tienen el mismo origen.

En el conjunto $\mathcal{F}$ de todos los vectores fijos del espacio ordinario se define la relación binaria de equivalencia $\sim$ siguiente:

$$\overrightarrow{PQ} \sim \overrightarrow{RS} \quad \Longleftrightarrow \quad PQSR \text{ es un paralogramo}$$



Dos vectores que son equivalentes en este sentido, se dice también que son equipolentes. A cada clase de equivalencia se le denomina ***vector libre***. En consecuencia, un vector libre es un conjunto de vectores fijos, todos ellos equivalentes entre sí y se representa por $\vec{\boldsymbol{v}} = [\overrightarrow{PQ}]$, indicando que $\overrightarrow{PQ}$ es un representante de todos los vectores su clase. Al conjunto de los vectores libres del espacio ordinario lo denotamos por $\mathcal{V}$.

Todos los vectores fijos nulos forman el vector libre cero. Un representante es $\overrightarrow{PP}$ (para un punto P cualquiera) y se designa por $\vec{\boldsymbol{0}}$.

Dado un vector libre $\vec{\boldsymbol{v}} = [\overrightarrow{PQ}]$ y fijado un punto O, existe un único vector fijo de su clase con origen en O, esto es, existe un único punto L tal que $OLQP$ es un paralelogramo; se dice que $\overrightarrow{OL}$ es el vector posición del punto L respecto a O.

Ocurre, entonces, que cada punto del espacio ordinario es origen de un representante de un vector libre.

Fijado un punto O, podemos sumar vectores libres y multiplicarlos por un número real, sin mas que tomar su representante con origen en O. Es decir, podemos definir las leyes de composición siguientes:

$$\begin{array}{ccc} \mathcal{V} \times \mathcal{V} & \longrightarrow & \mathcal{V} \\ (\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}) & \longmapsto & \vec{\boldsymbol{u}} + \vec{\boldsymbol{v}} \end{array} \qquad\qquad \begin{array}{ccc} \mathbb{R} \times \mathcal{V} & \longrightarrow & \mathcal{V} \\ (\lambda, \vec{\boldsymbol{u}}) & \longmapsto & \lambda\vec{\boldsymbol{u}} \end{array}$$

Tomando puntos P y Q, tales que $\vec{\boldsymbol{u}} = [\overrightarrow{OP}]$ y $\vec{\boldsymbol{v}} = [\overrightarrow{OQ}]$ y hacer la suma de vectores fijos: $\overrightarrow{OP} + \overrightarrow{OQ}$, que nos da un representante de la clase de la suma de los vectores libres y que denotamos por $\vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}$. Así mismo, $\lambda\overrightarrow{OP}$ es el presentante del vector libre producto del numero real λ y $\vec{\boldsymbol{u}}$, que se denota por $\lambda\vec{\boldsymbol{u}}$.

Las operaciones introducidas tienen, entre otras, las propiedades siguientes:

Propiedad asociativa: $\forall \vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \in \mathcal{V} : \vec{\boldsymbol{u}} + (\vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}) = (\vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}) + \vec{\boldsymbol{w}}$.
Propiedad conmutativa: $\forall \vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \in \mathcal{V} : \vec{\boldsymbol{u}} + \vec{\boldsymbol{v}} = \vec{\boldsymbol{v}} + \vec{\boldsymbol{u}}$.
Elemento neutro: $\forall \vec{\boldsymbol{u}} \in \mathcal{V}, \quad \dot{\exists}\, \vec{\boldsymbol{o}} \in \mathcal{V} : \quad \vec{\boldsymbol{o}} + \vec{\boldsymbol{u}} = \vec{\boldsymbol{u}} + \vec{\boldsymbol{o}} = \vec{\boldsymbol{u}}$.
Elemento opuesto: $\forall \vec{\boldsymbol{u}} = [\overrightarrow{PQ}] \in \mathcal{V}, \ \exists - \vec{\boldsymbol{u}} = [\overrightarrow{QP}] \in \mathcal{V} : \ -\vec{\boldsymbol{u}} + \vec{\boldsymbol{u}} = \vec{\boldsymbol{u}} + (-\vec{\boldsymbol{u}}) = \vec{\boldsymbol{o}}$.

Con estas propiedades, el conjunto de vectores libres $\mathcal{V}$, respecto de la operación suma, tiene estructura de grupo conmutativo: $(\mathcal{V}, +)$.

El producto de un número real por un vector libre goza de las siguientes propiedades:

Distributiva respecto de la suma de vectores: $\lambda(\vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}) = \lambda\vec{\boldsymbol{u}} + \lambda\vec{\boldsymbol{v}}$.
Distributiva respecto de la suma de escalares: $(\lambda + \mu)\vec{\boldsymbol{u}} = \lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{u}}$.
Asociatividad mixta: $\lambda(\mu\,\vec{\boldsymbol{u}}) = (\lambda\,\mu)\vec{\boldsymbol{u}}$.
El elemento unidad de $\mathbb{R}$ también es elemento unidad para esta operación: $1\,\vec{\boldsymbol{u}} = \vec{\boldsymbol{u}}$.

Estas propiedades que verifican las operaciones hechas con vectores libres y números reales nos sirven de justificación para, por abstracción, dar el concepto de espacio vectorial sobre un conjunto y un cuerpo de escalares más generales.

1.1 Definición.- *Se dice que un conjunto $\boldsymbol{E}$, no vacío, es un* ***espacio vectorial*** *sobre un cuerpo (conmutativo) K, si en $\boldsymbol{E}$ hay definida una operación interna (suma) $+\colon \boldsymbol{E} \times \boldsymbol{E} \longrightarrow \boldsymbol{E}$, $(\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}) \mapsto \vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}$, que lo dota de estructura de grupo conmutativo, es decir, se verifica, $\forall\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \in \boldsymbol{E}$:*

1) Propiedad asociativa: $\vec{\boldsymbol{u}} + (\vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}) = (\vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}) + \vec{\boldsymbol{w}}$.

2) Propiedad conmutativa: $\vec{\boldsymbol{u}} + \vec{\boldsymbol{v}} = \vec{\boldsymbol{v}} + \vec{\boldsymbol{u}}$.

3) Elemento neutro: $\forall\vec{\boldsymbol{u}} \in \boldsymbol{E}, \quad \exists\vec{\boldsymbol{o}} \in \boldsymbol{E} : \quad \vec{\boldsymbol{o}} + \vec{\boldsymbol{u}} = \vec{\boldsymbol{u}} + \vec{\boldsymbol{o}} = \vec{\boldsymbol{u}}$.

4) Elemento opuesto: $\forall\vec{\boldsymbol{u}} \in \boldsymbol{E}, \exists -\vec{\boldsymbol{u}} \in \boldsymbol{E} : -\vec{\boldsymbol{u}} + \vec{\boldsymbol{u}} = \vec{\boldsymbol{u}} + (-\vec{\boldsymbol{u}}) = \vec{\boldsymbol{o}}$.

En $\boldsymbol{E}$ hay definida una operación externa (producto) $\cdot\colon K \times \boldsymbol{E} \longrightarrow \boldsymbol{E}$, $(\lambda, \vec{\boldsymbol{v}}) \mapsto \lambda\vec{\boldsymbol{v}}$, verificando, $\forall\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \in \boldsymbol{E}$ y $\forall\lambda, \mu \in K$:

5) Distributiva respecto de la suma de vectores: $\lambda(\vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}) = \lambda\vec{\boldsymbol{u}} + \lambda\vec{\boldsymbol{v}}$.

6) Distributiva respecto de la suma de escalares: $(\lambda + \mu)\vec{\boldsymbol{u}} = \lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{u}}$.

7) Asociatividad mixta: $\lambda(\mu\,\vec{\boldsymbol{u}}) = (\lambda\,\mu)\vec{\boldsymbol{u}}$.

8) El elemento unidad 1_K de K también es elemento unidad para esta operación: $1_K\,\vec{\boldsymbol{u}} = \vec{\boldsymbol{u}}$.

Generalmente a los elementos de $\boldsymbol{E}$ se les llama vectores y a los de K escalares.
De los axiomas que figuran en esta definición de espacio vectorial, surgen una serie de propiedades de las que enunciamos algunas y de las que seguramente haremos uso posterior a lo largo de esta exposición, aunque no hagamos mención explícita a ellas.

1.2 Proposición.- *En un espacio vectorial $\boldsymbol{E}$ sobre el cuerpo K, se tiene:*

1) El elemento neutro es único (denotado por $\vec{\boldsymbol{o}}$).

2) El elemento opuesto de un vector es único.

3) $\lambda\vec{\boldsymbol{o}} = \vec{\boldsymbol{o}}, \quad \forall\lambda \in K$.

4) $0_K\vec{\boldsymbol{u}} = \vec{\boldsymbol{o}}, \quad \forall\vec{\boldsymbol{u}} \in \boldsymbol{E}$.

5) $(-\lambda)\vec{\boldsymbol{u}} = -(\lambda\vec{\boldsymbol{u}}), \quad \forall\lambda \in K, \quad \forall\vec{\boldsymbol{u}} \in E.$

6) $\forall\vec{\boldsymbol{u}} \in E, \forall\lambda \in K,$ *si* $\lambda\vec{\boldsymbol{u}} = \vec{\boldsymbol{o}} \Rightarrow (\lambda = 0_K$ *ó* $\vec{\boldsymbol{u}} = \vec{\boldsymbol{o}}).$

7) $\forall\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \in E, \forall\lambda \in K - \{0_K\},$ *si* $\lambda\vec{\boldsymbol{u}} = \lambda\vec{\boldsymbol{v}} \Rightarrow \vec{\boldsymbol{u}} = \vec{\boldsymbol{v}}.$

8) $\forall\vec{\boldsymbol{u}} \in E - \{\vec{o}\}, \forall\lambda, \mu \in K,$ *si* $\lambda\vec{\boldsymbol{u}} = \mu\vec{\boldsymbol{u}} \Rightarrow \lambda = \mu.$

9) $\forall\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \in E,$ *si* $\vec{\boldsymbol{u}} + \vec{\boldsymbol{v}} = \vec{\boldsymbol{u}} + \vec{\boldsymbol{w}} \Rightarrow \vec{\boldsymbol{u}} = \vec{\boldsymbol{v}}.$

1.3 Ejemplo.- *Sea K un cuerpo y*
$$K^n = \{(\alpha_1, \ldots, \alpha_n)\ /\alpha_i \in K\ (i = 1, \ldots n)\}$$
el producto cartesiano, $K \times \overset{n)}{\cdots} \times K$. K^n es un espacio vectorial $(K^n, +, \cdot)$ sobre K, con las operaciones:

$$(\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n),$$
$$\lambda(\alpha_1, \ldots, \alpha_n) = (\lambda\alpha_1, \ldots, \lambda\alpha_n),\ \lambda \in K.$$

1.4 Ejemplo.- *Sea $E = \left\{f / f\colon \mathbb{R} \to \mathbb{R},\ \text{una aplicación}\right\}$ y las operaciones:*

$$(f + g)(x) = f(x) + g(x), \qquad (\lambda f)(x) = \lambda f(x), \quad \forall\lambda, x \in \mathbb{R}.$$

E es un espacio vectorial sobre $\mathbb{R}$.

1.5 Ejemplo.- *Sea $\mathcal{P}_n[X] = \left\{p(X) \Big/ p(X) = \sum\limits_{k=0}^{n} p_k X^k, p_k \in \mathbb{R}\right\}$ el conjunto de los polinomios de grado $n \in \mathbb{N}$ con coeficientes reales en la indeterminada X $(p_0X^0 := p_0)$; en él definimos las operaciones:*

$$p(X) + q(X) = \sum_{k=0}^{n} p_k X^k + \sum_{k=0}^{n} q_k X^k = \sum_{k=0}^{n} (p_k + q_k) X^k,$$
$$\lambda\, p(X) = \lambda \left(\sum_{k=0}^{n} p_k X^k\right) = \sum_{k=0}^{n} (\lambda p_k) X^k.$$

$\mathcal{P}_n[X]$ es un espacio vectorial sobre $\mathbb{R}$.

1.6 Ejemplo.- *El conjunto de la matrices cuadradas de orden 2 con coeficientes reales:*

$$\mathcal{M}_2(\mathbb{R}) = \left\{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Big/ a_{ij} \in \mathbb{R}, (i, j = 1, 2)\right\},$$

con las operaciones:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix},$$
$$\lambda \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{pmatrix}, \quad \lambda \in \mathbb{R},$$

es un espacio vectorial sobre $\mathbb{R}$.

1.7 Definición.- *Un* `subespacio vectorial` *de un espacio vectorial* $\boldsymbol{E}$ *es un subconjunto* $\boldsymbol{F}\dot{\subset}\boldsymbol{E}$ *tal que con las mismas operaciones de* $\boldsymbol{E}$ *tiene estructura de espacio vectorial.*

El siguiente resultado da una caracterización para que un subconjunto de un espacio vectorial sea así mismo espacio vectorial.

1.8 Proposición.- *Un subconjunto* $\boldsymbol{F}$ *de un espacio vectorial* $\boldsymbol{E}$*, sobre* K*, es un subespacio vectorial si y sólo si*

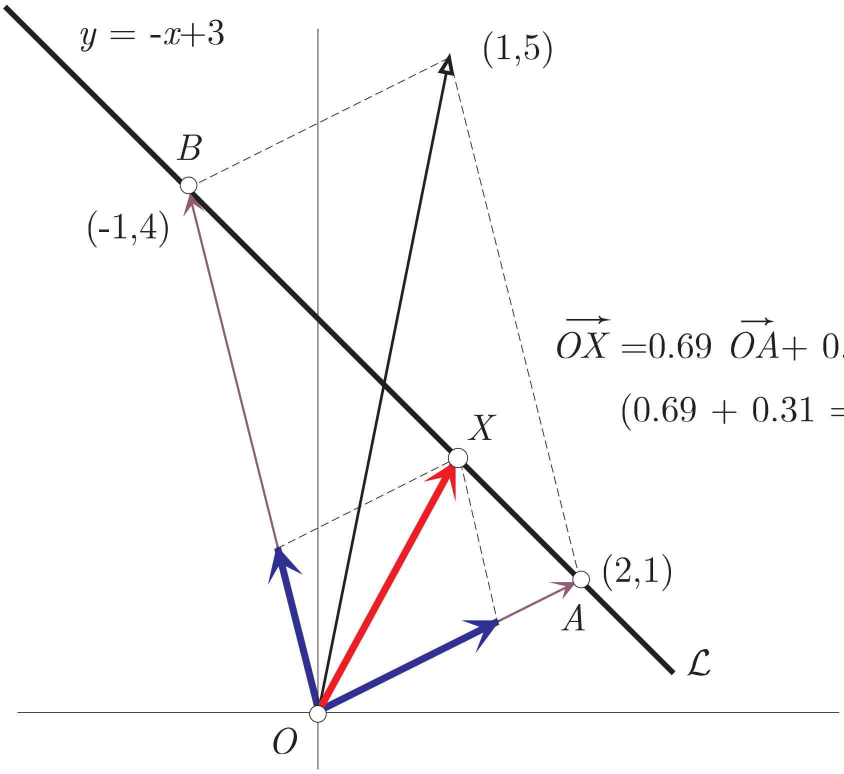
$$\forall \lambda,\mu \in K, \forall \vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \in \boldsymbol{F} \Rightarrow \lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}} \in \boldsymbol{F}.$$

1.9 Ejemplo.- *En el espacio vectorial* $(\mathbb{R}^2, +, \cdot)$*, el subconjunto*

$$\boldsymbol{F} = \{(x,y) \in \mathbb{R}^2 \ /y = mx\} = \{(x,mx) \ /x \in \mathbb{R} \ (m \in \mathbb{R})\} \subset \mathbb{R}^2$$

(recta que pasa por el origen) es un subespacio vectorial.

Pues, si $\lambda, \mu, x, x' \in \mathbb{R}$, $\lambda(x,mx) + \mu(x',mx') = (\lambda x, \lambda mx) + (\mu x', \mu mx') = (\lambda x + \mu x', m(\lambda x + \mu x')) \in \boldsymbol{F}$.



Applet-Cabri http://webpages.ull.es/users/amontes/cabri/g1-03.htm

1.10 Nota.- Sin embargo, el subconjunto (recta que no contiene al $(0,0)$),

$$\mathfrak{L} = \{(x, -x+3) \in \mathbb{R}^2 \ /x \in \mathbb{R}\},$$

no es un subespacio vectorial de $(\mathbb{R}^2, +, \cdot)$, ya que si $A(a, -a+3)$ y $B(b, -b+3)$ son dos elementos de $\mathfrak{L}$, y λ, μ son dos números reales arbitrarios:

$$\lambda(a, -a+3) + \mu(b, -b+3) = (\lambda a, \lambda(-a+3)) + (\mu b, \mu(-b+3)) =$$

$$= (\lambda a + \mu b, -(\lambda a + \mu b) + 3(\lambda + \mu)) \notin \mathfrak{L}.$$

No obstante, si tomamos sólo valores de λ y μ que cumplan $\lambda+\mu=1$, $\lambda(a,-a+3)+\mu(b,-b+3)\in\mathfrak{L}$. Así, podemos obtener todos los puntos de la recta $\mathfrak{L}$, tomando λ y μ, con la condición $\lambda+\mu=1$; es decir, a todo punto de la recta $\mathfrak{L}$, le podemos asociar un par de números reales (λ,μ), verificando $\lambda+\mu=1$. Ésta son las coordenadas baricéntricas (§3.2) de los puntos de $\mathfrak{L}$, respecto a $\{A,B\}$; en particular, $(1,0)$ son las coordenadas baricéntricas del punto A, y $(0,1)$ son las de B.

1.2 Dependencia e independencia lineal. Bases

1.11 Definición.- *Dado un subconjunto $\{\vec{u}_1,\dots,\vec{u}_r\}$ de r vectores de un espacio vectorial E, se dice que $\vec{u}\in E$ es* ***combinación lineal*** *de los r vectores dados si existen $\lambda_1,\dots,\lambda_r\in K$ (no todos nulos), tal que*

$$\vec{u}=\lambda_1\vec{u}_1+\cdots+\lambda_r\vec{u}_r.$$

1.12 Proposición.- *El conjunto de las combinaciones lineales de los elementos de $\mathsf{A}=\{\vec{u}_1,\dots,\vec{u}_r\}\subset E$ es un subespacio vectorial $\widetilde{\mathsf{A}}\dot{\subset}E$, denominado* ***subespacio vectorial generado por*** *A. $\widetilde{\mathsf{A}}$ es el menor espacio vectorial que contiene a A.*

1.13 Definición.- *Se dice que los vectores $\{\vec{u}_1,\dots,\vec{u}_r\}\subset E$ son* ***linealmente independientes*** *si*

$$\lambda_1\vec{u}_1+\cdots+\lambda_r\vec{u}_r=\vec{o}\Rightarrow\lambda_1=\cdots=\lambda_r=0_K.$$

Se dice que los vectores $\{\vec{u}_1,\dots,\vec{u}_r\}\subset E$ son ***linealmente dependientes*** *si existen escalares, no todos nulos, tales que*

$$\lambda_1\vec{u}_1+\cdots+\lambda_r\vec{u}_r=\vec{o}.$$

1.14 Definición.- *Un espacio vectorial se dice que es de* ***dimensión finita***, *si existe un subconjunto finito de generadores de él.*

1.15 Definición.- *Se llama* ***base*** *de un espacio vectorial a un subconjunto de vectores generadores de E y linealmente independientes.*

1.16 Proposición.- *En un espacio vectorial de dimension finita todas kas bases tienen el mismo número de elementos.*

1.17 Definición.- *Se llama **dimensión** de un espacio vectorial al número de elementos de una de sus bases. Si la dimensión es n, se usa $\dim_K \boldsymbol{E} = \dim \boldsymbol{E} = n$.*

Dada una base $\mathcal{B} = \{\vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$ de un espacio vectorial sobre K, para cualquier $\vec{\boldsymbol{u}} \in \boldsymbol{E}$ existen n únicos escalares $\lambda_1, \dots \lambda_n \in K$, que se llaman componentes de $\vec{\boldsymbol{u}}$ en la base $\mathcal{B}$, tales que:

$$\vec{\boldsymbol{u}} = \lambda_1 \vec{\boldsymbol{e}}_1 + \cdots + \lambda_r \vec{\boldsymbol{e}}_n.$$

1.18 Ejemplo.- *El conjunto $\boldsymbol{E} = \{(x, y, z) \in \mathbb{R}^3 \ / x + 2y - z = 0\}$, o sea, $\boldsymbol{E} = \{(x, y, x + 2y) \in \mathbb{R}^3 \ / x \in \mathbb{R}, y \in \mathbb{R}\}$, es un espacio vectorial de dimensión 2 y $\{\vec{\boldsymbol{u}}_1 = (-1, 1, 1), \vec{\boldsymbol{u}}_2 = (2, -1, 0)\}$ es una base.*

Los vectores $\vec{\boldsymbol{u}}_1$ y $\vec{\boldsymbol{u}}_2$ son elementos de $\boldsymbol{E}$ y linealmente independientes, pues:

$$\lambda(-1, 1, 1) + \mu(2, -1, 0) = (0, 0, 0) \Rightarrow \lambda = 0, \mu = 0.$$

Y si $\vec{\boldsymbol{v}} = (x, y, z) \in \boldsymbol{E}$, entonces existen $x + 2y \in \mathbb{R}$ y $x + y \in \mathbb{R}$ tales que $(x + 2y) \cdot (-1, 1, 1) + (x + y) \cdot (2, -1, 0) = \vec{\boldsymbol{v}}$.

1.19 Ejemplo.- *En el espacio vectorial $\mathcal{P}_2[X]$ de los polinomios de grado 2 con coeficientes reales (ver Ejemplo 1.5) los tres polinomios $p(X) = 2X^2 + X + 1$, $q(X) = 2X^2 + X$ y $r(X) = 2X^2 + 2$ forman una base.*

Son independientes: si $\lambda\, p(X) + \mu\, q(X) + \nu\, r(X) = 0 = 0_{\mathcal{P}_2[X]}$, es decir, si $(2\lambda + 2\mu + 2\nu)X^2 + (\lambda + \mu)X + (\lambda + 2\nu) = 0$. Entonces, $2\lambda + 2\mu + 2\nu = 0$, $\lambda + \mu = 0$, $\lambda + 2\nu = 0$, que tiene como única solución $\lambda = \mu = \nu = 0$.

Dado un polinomio cualquiera $aX^2 + bX + c \in \mathcal{P}_2[X]$, el sistema de ecuaciones $2\lambda + 2\mu + 2\nu = a$, $\lambda + \mu = b$, $\lambda + 2\nu = c$ tiene solución: $\lambda = -a + 2b + c$, $\mu = a - b - c$, $\nu = a/2 - b$.

Así, $\{p(X), q(X), r(X)\}$ es una base de $\mathcal{P}_2[X]$ y las componentes del polinomio $aX^2 + bX + c$ son $(-a + 2b + c, a - b - c, a/2 - b)$.

Cambio de base

Consideremos dos bases en $\boldsymbol{E}$: $\mathcal{B} = \{\vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$ y $\mathcal{B}' = \{\vec{\boldsymbol{e}}_1', \dots, \vec{\boldsymbol{e}}_n'\}$. La matriz cambio de bases, $M = (a_i^j)$, de $\mathcal{B}$ a $\mathcal{B}'$ viene dada por

$$\vec{\boldsymbol{e}}_i' = \sum_{j=1}^{n} a_i^j \vec{\boldsymbol{e}}_j \qquad (i = 1, \dots, n).$$

Si $\vec{\boldsymbol{u}} \in \boldsymbol{E} - \{\vec{0}\}$ se expresará por:

$$\vec{\boldsymbol{u}} = \sum_{i=1}^{n} x'^i \vec{\boldsymbol{e}}_i' = \sum_{i=1}^{n} x'^i \left(\sum_{j=1}^{n} a_i^j \vec{\boldsymbol{e}}_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_i^j x'^i \right) \vec{\boldsymbol{e}}_j = \sum_{j=1}^{n} x^j \vec{\boldsymbol{e}}_j,$$

luego

$$x^j = \sum_{i=1}^{n} a_i^j x'^i \qquad (j = 1, \dots, n).$$

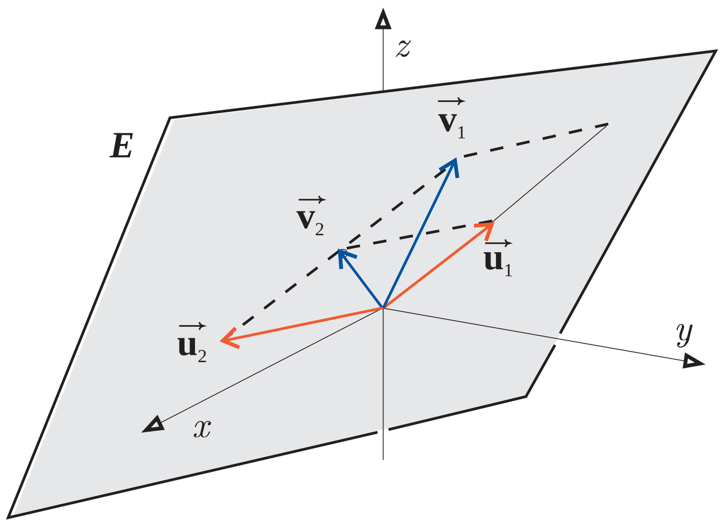
En notación matricial la relación entre las componentes del vector $\vec{\boldsymbol{u}}$, respecto a las dos bases, se expresa por:

$$\begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \cdots & a_n^n \end{pmatrix} \begin{pmatrix} x'^1 \\ x'^2 \\ \vdots \\ x'^n \end{pmatrix}.$$

Las ecuaciones obtenidas nos dan las componentes $(x^1, \ldots, x^n)$ de un vector $\vec{\boldsymbol{u}}$ respecto a la base $\mathcal{B}$, conocidas sus componentes $(x'^1, \ldots, x'^n)$ respecto a la base $\mathcal{B}$' y las componentes $(a_i^1, \ldots, a_i^n)$ de cada vector $\vec{\boldsymbol{e}}_i$' $(i = 1, \ldots, n)$ de la base $\mathcal{B}$' respecto a los de la base $\mathcal{B}$.

En notación simbólica, podemos poner las ecuaciones cambio de bases por $X = MX'$, siendo X y X' matrices columna formadas por las componentes $(x^1, \ldots, x^n)$ y $(x'^1, \ldots, x'^n)$ de un vector $\vec{\boldsymbol{u}}$ respecto a las bases $\mathcal{B}$ y $\mathcal{B}$', respectivamente, y $M = (a_i^j)$ la matriz cambio de bases, cuyas columnas son las componentes de los $\vec{\boldsymbol{e}}_i$' respecto a la base $\mathcal{B}$.

Para obtener las componentes $(x'^1, \ldots, x'^n)$ conocidas las $(x^1, \ldots, x^n)$ sólo tendremos que determinar la matriz inversa de M, es decir, $X' = M^{-1}X$.



1.20 Ejemplo.- *Sea $\{\vec{\boldsymbol{v}}_1 = (0,1,2), \vec{\boldsymbol{v}}_2 = (1,0,1)\}$ otra base del espacio vectorial **E** del* Ejemplo 1.18*, la matriz cambio de base que permite obtener las componentes de un vector respecto a la base $\{\vec{\boldsymbol{u}}_1, \vec{\boldsymbol{u}}_2\}$, conocidas sus componentes en la base $\{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2\}$, es* $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

Ya que, $\vec{\boldsymbol{v}}_1 = (0,1,2) = 2 \cdot (-1,1,1) + 1 \cdot (2,-1,0) = 2\vec{\boldsymbol{u}}_1 + 1\vec{\boldsymbol{u}}_2$ y $\vec{\boldsymbol{v}}_2 = (1,0,1) = 1 \cdot (-1,1,1) + 1 \cdot (2,-1,0) = 1\vec{\boldsymbol{u}}_1 + 1\vec{\boldsymbol{u}}_2$.

1.21 Ejemplo.- *La base canónica en el espacio vectorial $\mathcal{P}_2[X]$ de los polinomios de grado 2 con coeficientes reales es $\{1, X, X^2\}$, entonces la matriz de cambio de base que da las componentes de estos polinomios respecto la base dada en el* Ejemplo 1.19 *es:*

$$\begin{pmatrix} 1 & 2 & -1 \\ -1 & -1 & 1 \\ 0 & -1 & \frac{1}{2} \end{pmatrix}.$$

1.3 Intersección y suma de subespacios vectoriales

1.22 Proposición.- *Si F y G son subespacios vectoriales de un espacio vectorial E, la intersección $F \cap G$ es un subespacio vectorial de E.*

1.23 Definición.- *Sean F y G dos subespacios vectoriales de un espacio vectorial E, se llama **suma** de F y G al subespacio vectorial*

$$F + G = \{\vec{u} + \vec{v} \,/\, \vec{u} \in F, \vec{v} \in G\}$$

1.24 Proposición.- *La suma $F + G$ de dos subespacios vectoriales de E es el menor subespacio, $\widetilde{F \cup G}$, que contiene a $F \cup G$.*

1.25 Definición.- *Sean E un espacio vectorial y F, G y H subespacios vectoriales de E, se dice que H es **suma directa** de F y G, y se escribe $F \oplus G = H$, si y sólo si $F + G = H$ y $F \cap G = \{\vec{0}\}$.*

1.26 Definición.- *Sea E un espacio vectorial y sean F y G dos subespacios vectoriales de E, se dice que F y G son **suplementarios** si $F + G = E$ y $F \cap G = \{\vec{0}\}$.*

1.27 Proposición.- *Sean F y G dos subespacios vectoriales de un espacio vectorial E de dimensión finita. Entonces se verifica la siguiente relación entre las dimensiones:*

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G).$$

1.28 Ejemplo.- *Si F y G son subespacios vectoriales suplementarios de un espacio vectorial E, $\dim E = \dim F + \dim G$.*

1.29 Ejemplo.- *Sean los subespacios vectoriales (planos que pasan por el origen) de $\mathbb{R}^3$:*

$$H_1 = \{(x, y, z) \in \mathbb{R}^3 \,/ z = 0\} \text{ y } H_2 = \{(x, y, z) \in \mathbb{R}^3 \,/ x = y\};$$

su intersección es la recta $L = \{(x, y, z) \in \mathbb{R}^3 \,/ z = 0, x = y\}$. Entonces:

$$\dim(H_1 + H_2) = \dim H_1 + \dim H_2 - \dim L = 2 + 2 - 1 = 3 = \dim \mathbb{R}^3.$$

1.30 Nota.- Sea F un subespacio vectorial de E; se demuestra que F admite por lo menos un subespacio G tal que F y G son suplementarios. Existen en general diversos suplementarios de F. La existencia de suplementarios, para espacios de dimensión finita, se puede demostrar usando el teorema de la base incompleta.
Para la demostración en general [(1)] se hace uso del axioma de elección de Zorn.

(1) [17, Ejercicio I.1] (Existencia de subespacios suplementarios)

1.4 Aplicaciones lineales

1.31 Definición.- *Sea* E *y* F *dos espacios vectoriales sobre un mismo cuerpo* K*. Una aplicación* $f: E \to F$ *se dice que es* **lineal** *si cumple:*

$$f(\vec{u}+\vec{v}) = f(\vec{u}) + f(\vec{v}), \quad f(\lambda\vec{u}) = \lambda f(\vec{u}), \quad \forall \vec{u}, \vec{v} \in E, \forall\lambda \in K.$$

1.32 Proposición.- *La condición necesaria y suficiente para que una aplicación* $f: E \to F$ *entre espacios vectoriales sea lineal es que:*

$$f(\lambda\vec{u}+\mu\vec{v}) = \lambda f(\vec{u}) + \mu f(\vec{v}), \quad \forall\vec{u}, \vec{v} \in E, \forall\lambda, \mu \in K.$$

Al conjunto de las aplicaciones lineales de E en F lo denotamos por $\mathcal{L}_K(E, F)$.

1.33 Definición.- *Sea* $f \in \mathcal{L}_K(E, F)$*, se denomina* **imagen** *de* f *al subconjunto* $\mathrm{Im}(f) = f(E) = \{y \in F \,/ \exists\vec{u} \in E, \vec{v} = f(\vec{u})\} \dot{\subset} F$ *y* **núcleo** *de* f *al subconjunto* $\mathrm{Ker}(f) = \{\vec{u} \in E \,/ f(\vec{u}) = \vec{o}_F\} = f^{-1}(\vec{o}_F) \dot{\subset} E$*.*

1.34 Proposición.- *Sea* $f: E \to E'$ *una aplicación lineal, se verifica que:*

- $f(\vec{o}_E) = \vec{o}_{E'}$.
- *Si* $F \dot{\subset} E$ *es un subespacio vectorial* $\Rightarrow f(F) \dot{\subset} E'$ *es subespacio vectorial.*
- *Si* $F' \dot{\subset} E'$ *es un subespacio vectorial* $\Rightarrow f^{-1}(F') \dot{\subset} E$ *es subespacio vectorial.*
- *Si* $\{\vec{u}_1, \ldots, \vec{u}_r\} \subset E$ *es un sistema generador de* $F \dot{\subset} E \Rightarrow$ $\{f(\vec{u}_1), \ldots, f(\vec{u}_r)\} \subset E'$ *es un sistema generador de* $f(F) \dot{\subset} E'$*.*

1.35 Proposición.- *Sea* $f \in \mathcal{L}_K(E, F)$ *entonces el núcleo* $\mathrm{Ker}(f)$ *y la imagen* $\mathrm{Im}(f) = f(E)$ *son subespacios vectoriales de* E *y* F*, respectivamente.*

Sea E un espacio vectorial sobre un cuerpo conmutativo K y X un subespacio vectorial de E distinto de E. Sea $\mathcal{F}$ el conjunto de los subespacios vectoriales L de E tales que $L \cap X = \{0\}$. Se considera en $\mathcal{F}$ la relación de orden parcial definida por la inclusión $L \subset M$.

a) Demostrar que toda parte totalmente ordenada de $\mathcal{F}$ admite una cota superior (véase tomo I, pág. 33).

b) Demostrar que todo elemento maximal de $\mathcal{F}$ es un subespacio suplementario de X (véase tomo I, pág. 35).

c) La existencia de un subespacio suplementario de X resulta entonces del axioma de Zorn: si $\mathcal{F}$ es un conjunto ordenado tal que toda parte totalmente ordenada de $\mathcal{F}$ admite una cota superior, entonces $\mathcal{F}$ admite por lo menos un elemento maximal.

Ver también:
Teorema 3.5.6, pág 94 en:
Daniel Hernández Ruipérez.- Álgebra lineal. Colección: MU (Manuales Universitarios), 25. Universidad de Salamanca. 1985.

1.36 Proposición.- *Sea $f \in \mathcal{L}_K(E, F)$:*
f *es inyectiva* $\Leftrightarrow \operatorname{Ker}(f) = \{\vec{o}_E\}$.
f *es suprayectiva* $\Leftrightarrow \operatorname{Im}(f) = F$.
f *biyectiva* $\Leftrightarrow f^{-1} \in \mathcal{L}_K(F, E)$.

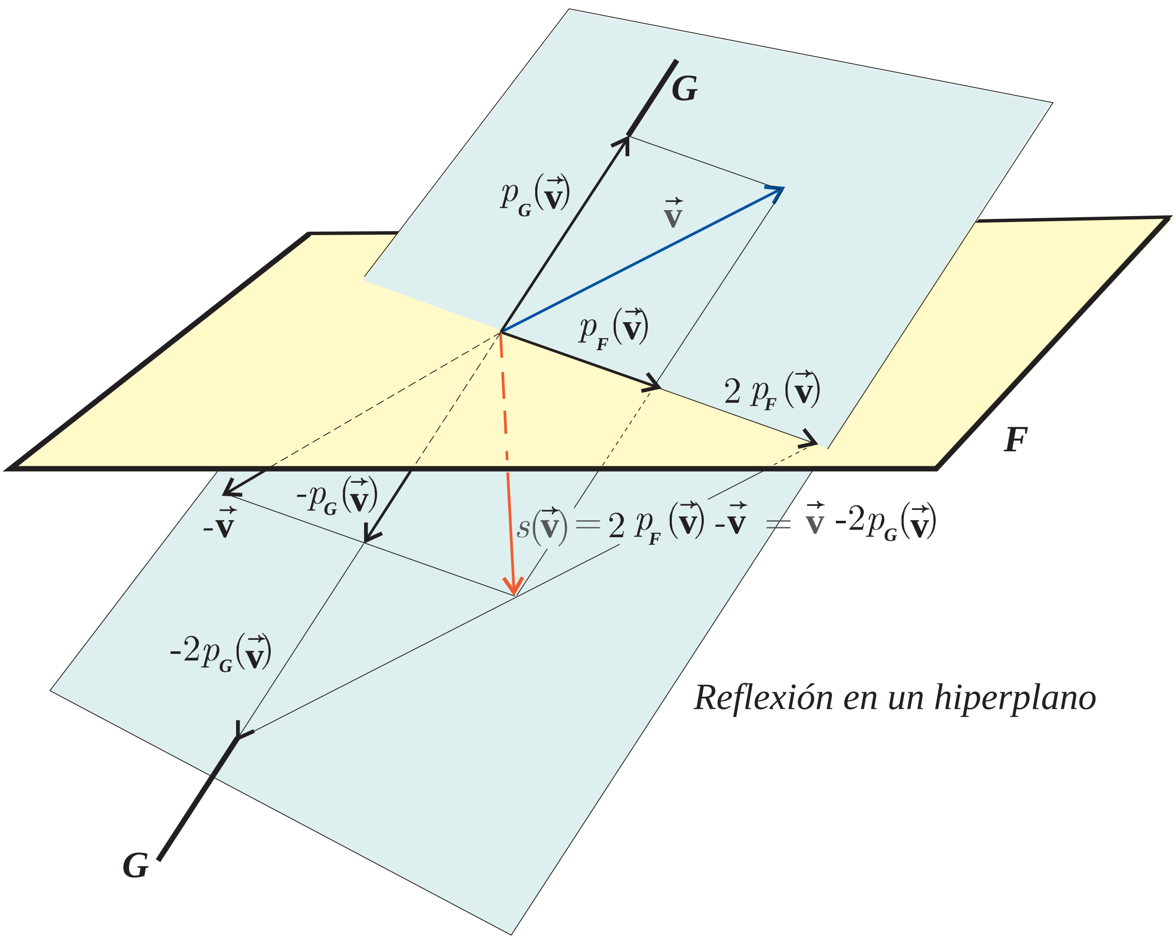
1.37 Proposición.- *Si E y F son espacios vectoriales de dimensión finita sobre un mismo cuerpo K y $f \in \mathcal{L}_K(E, F)$:*

$$\dim \operatorname{Ker}(f) + \dim \operatorname{Im}(f) = \dim E.$$

1.38 Definición.- *Una aplicación lineal biyectiva entre espacios vectoriales se dice que es un **isomorfismo**.*

1.39 Definición.- *Sean F y G dos subespacios vectoriales suplementarios de un espacio vectorial E, $F \oplus G = E$, entonces, para cualquier $\vec{v} \in E$, existe un único vector $p_F(\vec{v}) \in F$ y un único vector $p_G(\vec{v}) \in G$, tales que $\vec{v} = p_F(\vec{v}) + p_G(\vec{v})$.*

*Las aplicaciones $p_F: E \longrightarrow E$ y $p_G: E \longrightarrow E$ reciben respectivamente los nombres de **proyector** de E sobre F (paralelamente a G) y proyector de E sobre G (paralelamente a F).*



Reflexión en un hiperplano

Algunas propiedades de los proyectores, los que llamaremos proyectores lineales cuando tratemos proyecciones afines, debido a la primera de las propiedades que vamos a citar:

1) Todo proyector es una aplicación lineal.

2) Si p_F y p_G son proyectores de E sobre subespacios suplementarios F y G, respectivamente, entonces:

$$\mathrm{Im}(p_F) = F, \quad \mathrm{Im}(p_G) = G \qquad \mathrm{Ker}(p_F) = G, \quad \mathrm{Ker}(p_G) = F.$$

$$p_F \circ p_G = p_G \circ p_F = 0_E, \qquad p_F + p_G = 1_E.$$

3) Cualquier proyector verifica que $p \circ p = p$.

4) Si $p\colon E \to E$ es una aplicación lineal tal que $p^2 = p$, entonces es un proyector de E sobre $F = \mathrm{Im}(p)$ (paralelamente a $G = \mathrm{Ker}(p)$).

En efecto, veamos que $F = p(E)$ y $G = \mathrm{Ker}(p)$ son suplementarios.

— Si $\vec{w} \in p(E) \cap \mathrm{Ker}(p)$, entonces $\vec{w} = \vec{o}$. En efecto, $p(\vec{w}) = \vec{o}$ y existe $\vec{v} \in E$ tal que $p(\vec{v}) = \vec{w}$. Luego, $\vec{o} = p(\vec{w}) = p^2(\vec{v}) = p(\vec{v}) = \vec{w}$.

— Si $\vec{v} \in E$, lo podemos poner de la forma $\vec{v} = p(\vec{v}) + (\vec{v} - p(\vec{v})) \in p(E) + \mathrm{Ker}(p)$.

1.40 Definición.- *Sean F y G dos subespacios vectoriales suplementarios de E, la **simetría** (o **reflexión**) respecto a F y paralelamente a G es la aplicación lineal $s\colon E \to E$ definida, para todo $\vec{v} \in E$, por*

$$s(\vec{v}) = 2p_F(\vec{v}) - \vec{v}.$$

Debido a que $p_F + p_G = 1_E$, también se tiene que:

$$s(\vec{v}) = p_F(\vec{v}) - p_G(\vec{v}) \quad \text{y} \quad s(\vec{v}) = \vec{v} - 2p_G(\vec{v}).$$

Una simetría s verifica:

$$s^2 = 1_E, \quad s_{|F} = 1_F, \quad s_{|G} = -1_G.$$

Matriz asociada a una aplicación lineal

1.41 Proposición.- *Sean E y F espacios vectoriales de dimensiones n y m, respectivamente, y sean $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ y $\mathcal{B}' = \{\vec{e}\,'_1, \dots, \vec{e}\,'_m\}$ bases de E y F, respectivamente. Una aplicación lineal $f\colon E \to F$ se puede expresar por las siguientes ecuaciones:*

$$\begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}.$$

donde $(x^1, \dots, x^n)$ son las componentes de un vector $\vec{u}$ respecto de $\mathcal{B}$ y $(y^1, \dots, y^m)$ las componentes del vector $\vec{v} = f(\vec{u})$ respecto de la base $\mathcal{B}'$. Las columnas de la matriz (a_i^j) son las componentes del vector $f(\vec{e}_i) = \sum_{j=1}^{m} a_i^j \vec{e}\,'_j$ respecto a la base $\mathcal{B}'$, $(i = 1, \dots, n)$.

Con notación abreviada, la ecuación matricial de una aplicación lineal f se puede poner de la forma $Y = MX$, siendo X e Y las matrices columnas formadas por las componentes de los vectores $\vec{u}$ e $\vec{v}$, y $M = (a_i^j)$, denominada matriz asociada a f respecto de las bases $\mathcal{B}$ y $\mathcal{B}'$.

1.42 Ejemplo.- *Consideremos en un espacio vectorial E, de dimensión n, dos bases $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ y $\mathcal{B}' = \{\vec{e}_1', \ldots, \vec{e}_n'\}$ y sea $1_E\colon E \to E$ la aplicación identidad (que claramente es lineal), entonces la matriz asociada a 1_E respecto de las bases $\mathcal{B}$ y $\mathcal{B}'$ es la matriz cambio de bases de $\mathcal{B}'$ a $\mathcal{B}$* (ver pág. 7).

En efecto, para determinar las columnas de la matriz asociada a 1_E, debemos hallar las componentes de $1_E(\vec{e}_i) = \vec{e}_i$ en la base $\mathcal{B}' = \{\vec{e}_1', \ldots, \vec{e}_n'\}$.

1.43 Proposición.- *Sean E, F y G espacios vectoriales de dimensiones n, m y p, respectivamente, y sean $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}, \mathcal{B}' = \{\vec{e}_1', \ldots, \vec{e}_m'\}$ y $\mathcal{B}'' = \{\vec{e}_1'', \ldots, \vec{e}_p''\}$ bases de E, F y G, respectivamente. Sean las aplicaciones lineales $f\colon E \to F$ y $g\colon F \to G$, con matrices asociadas M y M', entonces la matriz asociada a la composición $g \circ f\colon E \to G$ tiene como matriz asociada, respecto a las bases $\mathcal{B}$ y $\mathcal{B}''$, la matriz producto $M'' = M'\,M$.*

En efecto, si $\vec{e}_i$ es de la base $\mathcal{B}$, $M = (a_i^k)_{\substack{i=1,\ldots n\\ k=1,\ldots m}}$ y $M' = (b_k^j)_{\substack{j=1,\ldots p\\ k=1,\ldots m}}$:

$$(g \circ f)(\vec{e}_i) = g(f(\vec{e}_i)) = g\Big(\sum_{k=1}^{m} a_i^k \vec{e}_k'\Big) = \sum_{k=1}^{m} a_i^k g(\vec{e}_k') = \sum_{k=1}^{m} a_i^k \Big(\sum_{j=1}^{p} b_k^j \vec{e}_j''\Big).$$

Así, la entrada de la fila j y columna k de la matriz asociada a $g \circ f$ es $c_i^j = \sum_{j=1}^{p} b_k^j a_i^k$, que es la correspondiente entrada del matrices producto $M'M$.

Cambios de base en una aplicación lineal

1.44 Proposición.- *Sean E y F dos espacio vectorial de dimension finita, $\mathcal{B}_E$ y $\mathcal{B}_E'$ dos bases de E con M_E la matriz cambio de bases de $\mathcal{B}_E$ a $\mathcal{B}_E'$* (pág. 7) *y $\mathcal{B}_F$ y $\mathcal{B}_F'$ dos bases de F con M_F matriz cambio de bases de $\mathcal{B}_F$ a $\mathcal{B}_F'$. Si $f\colon E \to F$ es una aplicación lineal con matriz M asociada a las bases $\mathcal{B}_E$ y $\mathcal{B}_F$ y con matriz M' asociada a las bases $\mathcal{B}_E'$ y $\mathcal{B}_F'$, la relación entre las matrices M y M' es $M' = M_F\, M\, M_E^{-1}$.*

Demostración.- Esta igualdad se obtiene sin mas que tener presente cual es la matriz asociada a las aplicaciones identidad de $1_E \in \mathcal{L}_K(E, E)$ y $1_F \in \mathcal{L}_K(F, F)$ (ver Ejemplo 1.42) y el siguiente diagrama conmutativo:

$$\begin{array}{ccc} E & \xrightarrow{\quad f \quad} & F \\ 1_E^{-1}\uparrow & & \downarrow 1_F \\ E & \xrightarrow{1_F^{-1} \circ f \circ 1_E} & F \end{array}$$

$\blacksquare$

1.45 Ejercicio.- Encontrar una aplicación lineal $f\colon \mathbb{R}^3 \to \mathbb{R}^2$, cuyo núcleo sea el subespacio $E = \{(x, y, z) \in \mathbb{R}^3 \;/ x + y + z = 0\}$ y cuya imagen sea $F = \{(x, x) \in \mathbb{R}^2 \;/ x \in \mathbb{R}\}$. ¿Cuál es su matriz asociada respecto a las bases canónica de $\mathbb{R}^3$ y $\mathbb{R}^2$?

Podemos tomar como base de $I\!R^3$, $\mathcal{B} = \{\vec{\boldsymbol{u}}_1 = (1,0,0), \vec{\boldsymbol{u}}_2 = (1,-1,0), \vec{\boldsymbol{u}}_3 = (1,0,-1)\}$, donde $\vec{\boldsymbol{u}}_2$ y $\vec{\boldsymbol{u}}_3$ están en $\boldsymbol{E}$ (núcleo de f); y como base de $I\!R^2$ la canónica, $\mathcal{B}'_0 = \{\vec{\boldsymbol{e}}_1{}' = (1,0), \vec{\boldsymbol{e}}_2{}' = (0,1)\}$. Como la imagen de todo vector, no contenido en $\boldsymbol{E}$, debe pertenecer a $\boldsymbol{F}$, se tiene, en particular, que $f(\vec{\boldsymbol{u}}_1) = (a,a)$, para $a \in I\!R$. Luego, la matriz asociada a f respecto a las bases $\mathcal{B}$ y $\mathcal{B}'_0$ es $M = \begin{pmatrix} a & 0 & 0 \\ a & 0 & 0 \end{pmatrix}$. Para determinar la matriz asociada a f, respecto a las bases canónicas de $I\!R^3$, $\mathcal{B}_0 = \{\vec{\boldsymbol{e}}_1 = (1,0,0), \vec{\boldsymbol{e}}_2 = (0,1,0), \vec{\boldsymbol{e}}_3 = (0,0,1)\}$, y $\mathcal{B}'_0$ de $I\!R^3$, observemos que la matriz cambio de bases de $\mathcal{B}$ a $\mathcal{B}'_0$ es $M_{I\!R^3} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, (¡¡que en este caso particular coincide con su inversa!!). Luego,

$$M' = M_{I\!R^2}\, M\, M_{I\!R^3}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ a & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} a & a & a \\ a & a & a \end{pmatrix}.$$

Para la aplicación f que hemos tomado, se verifica:

$$f((x,y,z)) = a(x+y+z)\,(1,1), \quad \forall (x,y,z) \in I\!R^3.$$

1.5 El espacio dual

Consideramos ahora un caso particular de aplicaciones lineales, que es cuando el espacio vectorial de llegada es K. En este caso a una aplicación lineal $f\colon \boldsymbol{E} \to K$ se le denomina ***forma lineal*** y al espacio vectorial

$$\boldsymbol{E}^* = \mathcal{L}_K(\boldsymbol{E}, K)$$

de las formas lineales sobre $\boldsymbol{E}$ le denominamos ***espacio dual*** de $\boldsymbol{E}$.

Si $\dim \boldsymbol{E} = n$, a una base $\mathcal{B} = \{\vec{\boldsymbol{v}}_1, \dots, \vec{\boldsymbol{v}}_n\}$ se le asocian las formas lineales

$$\theta^i\colon \boldsymbol{E} \to K, \qquad \theta^i(\vec{\boldsymbol{v}}_j) = \delta^i_j = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases},$$

que forman una base $\mathcal{B}^* = \{\theta^1, \dots, \theta^n\}$ de $\boldsymbol{E}^*$, denominada ***base dual*** de $\mathcal{B}$.

1.46 Ejemplo.- *En $I\!R^3$ se consideran las formas lineales*

$$\varphi_1(x,y,z) = 2x - y + 3z,$$
$$\varphi_2(x,y,z) = 3x - 3y + z,$$
$$\varphi_3(x,y,z) = 4x + 7y + z.$$

$\{\varphi_1, \varphi_2, \varphi_3\}$ es una base de $(I\!R^3)^*$.

Ya que si $\lambda\varphi_1 + \mu\varphi_2 + \nu\varphi_3 = 0_{(I\!R^3)^*}$ es decir, $\forall (x,y,z) \in I\!R^3$:

$$\lambda(2x - y + 3z) + \mu(3x - 3y + z) + \nu(4x + 7y + z) = 0.$$

$$2\lambda + 3\mu + 4\nu = 0, \quad -\lambda - 3\mu + 7\nu = 0, \quad 3\lambda + \mu + \nu = 0 \Rightarrow \lambda = \mu = \nu = 0.$$

Dadas una aplicación lineal $f\colon \boldsymbol{E} \to \boldsymbol{F}$ y una forma lineal $\omega \in \boldsymbol{F}^*$, la composición $\omega \circ f \in \boldsymbol{E}^*$. Definimos así la ***aplicación lineal traspuesta*** de f o ***aplicación dual*** de f:

$$^{\mathrm{t}}f\colon \boldsymbol{F}^* \to \boldsymbol{E}^*, \qquad \omega \in \boldsymbol{F}^* \mapsto {}^{\mathrm{t}}f(\omega) = \omega \circ f \in \boldsymbol{E}^*,$$

que es una aplicación lineal.

1.47 Proposición.- *Sean $f, g\colon \boldsymbol{E} \to \boldsymbol{F}$ aplicaciones lineales, entonces:*

$$(g \circ f)^* = f^* \circ g^*.$$

Demostración.-

$(g \circ f)^*(\omega) = \omega \circ (g \circ f) = (\omega \circ g) \circ f = f^*(\omega \circ g) = f^*(g^*(\omega)) = (f^* \circ g^*)(\omega).$ ◪

1.48 Proposición.- *Sea una aplicación lineal $f\colon \boldsymbol{E} \to \boldsymbol{F}$ que tiene a $A = (a_i^j)$ como matriz asociada a las bases $\mathcal{B}_E = \{\vec{\boldsymbol{v}}_1, \dots, \vec{\boldsymbol{v}}_n\}$ de $\boldsymbol{E}$ y $\mathcal{B}_F = \{\vec{\boldsymbol{w}}_1, \dots, \vec{\boldsymbol{w}}_n\}$ de $\boldsymbol{F}$. Su aplicación traspuesta $f^*\colon \boldsymbol{F}^* \to \boldsymbol{E}^*$ tiene a $B = (B_i^j)$ como matriz asociada a las bases duales $\mathcal{B}_F^* = \{\omega^1, \dots, \omega^n\}$ de $\boldsymbol{F}^*$ y $\mathcal{B}_E^* = \{\theta^1, \dots, \theta^n\}$ de $\boldsymbol{E}^*$. Entonces B es la matriz traspuesta de A ($B = {}^tA$).*

Demostración.- $f(\vec{\boldsymbol{v}}_i) = \sum_{k=1}^{n} a_i^k \vec{\boldsymbol{w}}_k$ y $f^*(\omega^j) = \sum_{k=1}^{n} b_k^j \theta^k$.

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix} \qquad B = \begin{pmatrix} b_1^1 & b_2^1 & \dots & b_1^n \\ b_2^1 & b_2^2 & \dots & b_2^n \\ \vdots & \vdots & \ddots & \vdots \\ b_n^1 & b_n^2 & \dots & b_n^n \end{pmatrix}$$

$$b_i^j = (f^*(\omega^j))(\vec{\boldsymbol{v}}_i) = \omega^j(f(\vec{\boldsymbol{v}}_i)) = \omega^j\left(\sum_{k=1}^{n} a_i^k \vec{\boldsymbol{w}}_k\right) = a_i^j.$$

◪

TEMA 2
Espacios afines

Geométricamente, las rectas, las curvas, planos y las superficies suelen ser considerados como conjuntos de puntos con algunas características especiales, que yacen en un espacio formado por "puntos". No obstante, es natural sentir inquietud por las propiedades geométricas invariantes por ciertas transformaciones, por ejemplo, traslaciones, rotaciones, proyecciones, etc. Se podría modelar el espacio de puntos como un espacio vectorial, pero esto no es muy satisfactorio para un buen número de razones. Una de las razones es que el punto correspondiente al vector cero ($\vec{o}$), llamado el origen, juega un papel especial, cuando en realidad no hay razón para tener un origen privilegiado. Otra razón es que ciertas nociones, como el paralelismo, no pueden ser tratadas. Pero la razón más profunda es que los espacios vectoriales y espacios afines realmente tienen diferentes geometrías. Las propiedades geométricas de un espacio vectorial son invariantes bajo el grupo de aplicaciones lineales biyectivas, mientras que las propiedades geométricas de un espacio afín son invariantes bajo el grupo de aplicaciones afines biyectivas (afinidades), y estos dos grupos no son isomorfos. Hablando sin mucho rigor, se puede decir que hay más aplicaciones afines que los aplicaciones lineales.

Los espacios afines proporcionan un buen marco para hacer geometría. En particular, es posible tratar a puntos, curvas, superficies, etc., de manera intrínseca, es decir, independientemente de cualquier elección concreta de un sistema de coordenadas. Al igual que en la física, esto es muy conveniente para entender realmente lo que está pasando. Por supuesto, los sistemas de coordenadas tiene que ser elegido para finalmente, realizar cálculos, pero uno debe aprender a resistir la tentación de recurrir a sistemas de coordenadas hasta que sea realmente necesario.

Los espacios afines son el marco adecuado para hacer frente a movimientos, trayectorias, y las fuerzas físicas, entre otras cosas. Por lo tanto, la geometría afín es crucial para mejorar la presentación de la cinemática, dinámica, y otras partes del física (por ejemplo, la elasticidad). Después de todo, un movimiento rígido es una aplicación afín, pero no una aplicación lineal en general. Además, dado una matriz A de orden $m \times n$ y un vector $b \in \mathbb{R}^n$, el conjunto $\mathcal{F} = \{x \in \mathbb{R}^n / Ax = b\}$ de las soluciones del sistema $Ax = b$ es un espacio afín (Ejemplo 2.8, Nota 3.5), pero no un espacio vectorial en general. ([4, pág. 6])

En esencia, el concepto de espacio afín es una estructura matemática que se ideó para "capturar" lo que en nuestra experiencia física y sensorial experimentamos como "Espacio".

2.1 Espacios afines. Definición y ejemplos

Definición de espacio afín

La propiedad fundamental de los vectores libres nos permite definir una aplicación biyectiva entre los puntos del espacio ordinario $\mathcal{E}$ y los vectores del espacio vectorial $\mathcal{V}$ de vectores libres, de la siguiente forma: Fijado un punto P arbitrario de $\mathcal{E}$, definimos una aplicación de $\mathcal{E}$ en $\mathcal{V}$ (biyectiva) que a cada punto $Q \in \mathcal{E}$ le hace corresponder el vector libre $\vec{\boldsymbol{v}} = [\overrightarrow{PQ}]$. O bien, dado un punto $P \in \mathcal{E}$ y un vector $\vec{\boldsymbol{v}} \in \mathcal{V}$, existe un único punto $Q \in \mathcal{E}$, extremo del vector fijo con origen en P y representante de la clase de $\vec{\boldsymbol{v}}$. Esto lo podemos formalizar definiendo una operación externa sobre $\mathcal{E}$, con dominio de operadores $\mathcal{V}$:

$$\phi\colon \mathcal{E} \times \mathcal{V} \longrightarrow \mathcal{E}, \qquad (P, \vec{\boldsymbol{v}}) \mapsto Q,$$

siendo Q el único punto del espacio ordinario cumpliendo que $\vec{\boldsymbol{v}} = [\overrightarrow{PQ}]$. A esta operación [(1)] la denotamos por "+":

$$\phi(P, \vec{\boldsymbol{v}}) = P + \vec{\boldsymbol{v}} = Q.$$

Que satisface, entre otras, las siguientes propiedades:

1) $\forall P, Q \in \mathcal{E},\ \dot{\exists} \vec{\boldsymbol{v}} \in \mathcal{V} :\ P + \vec{\boldsymbol{v}} = Q.$
2) $\forall P \in \mathcal{E}, \forall \vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \in \mathcal{V} : (P + \vec{\boldsymbol{u}}) + \vec{\boldsymbol{v}} = P + (\vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}).$

Estas consideraciones en el espacio ordinario nos sirve para dar una primera definición de espacio afín como sigue:

2.1 Definición.- *Dados un conjunto $\mathcal{A}$ no vacío (a cuyos elementos llamamos puntos) y un espacio vectorial $\boldsymbol{E}$ (a cuyos elementos llamamos vectores) sobre un cuerpo conmutativo K (a cuyos elementos llamamos escalares), diremos que $\mathcal{A}$ es un* ***espacio afín*** *asociado a $\boldsymbol{E}$ si se tiene definida una operación externa en $\mathcal{A}$ con dominio de operadores en $\boldsymbol{E}$:*

$$\phi\colon \mathcal{A} \times \boldsymbol{E} \to \mathcal{A}, \qquad (P, \vec{\boldsymbol{v}}) \longmapsto \phi(P, \vec{\boldsymbol{v}}) := P + \vec{\boldsymbol{v}}$$

que satisface a los siguientes axiomas:

A_1 $\forall P \in \mathcal{A}, \phi_P\colon \boldsymbol{E} \to \mathcal{A}$ definida por $\phi_P(\vec{\boldsymbol{v}}) = \phi(P, \vec{\boldsymbol{v}})$, es biyectiva; es decir: $\forall P, Q \in \mathcal{A},\ \dot{\exists} \vec{\boldsymbol{v}} \in \boldsymbol{E} :\ P + \vec{\boldsymbol{v}} = Q$.

A_2) Asociatividad de la ley externa respecto a la suma en $\boldsymbol{E}$.

$$\forall P \in \mathcal{A}, \forall \vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \in \boldsymbol{E} : \phi(P, \vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}) = \phi(\phi(P, \vec{\boldsymbol{u}}), \vec{\boldsymbol{v}});$$

es decir: $P + (\vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}) = (P + \vec{\boldsymbol{u}}) + \vec{\boldsymbol{v}}$.

La estructura de espacio afín descrita por esta definición la denotamos por $(\mathcal{A}, \boldsymbol{E}, \phi)$ o por $(\mathcal{A}, \boldsymbol{E}, +)$ y al vector único $\phi_P^{-1}(Q)$ asociado a dos puntos $P, Q \in \mathcal{A}$, por el Axioma A_1, lo denotamos por $\overrightarrow{PQ}$ y escribimos $P + \overrightarrow{PQ} = Q$.

Los axiomas de espacio afín los reformulamos de otras dos formas equivalentes, lo cual abordamos en las dos definiciones siguientes.

---

(1) Para no recargar el lenguaje, en vez usar la notación $\phi\Big((P, \vec{\boldsymbol{v}})\Big)$ para aplicar ϕ a un elemento que un par ordenados $(P, \vec{\boldsymbol{v}})$, que parece ser lo más correcto, usamos simplemente $\phi(P, \vec{\boldsymbol{v}})$.

2.1' Definición.- *Dados un conjunto $\mathcal{A}$ no vacío y un espacio vectorial $\boldsymbol{E}$ sobre un cuerpo conmutativo K, diremos que $\mathcal{A}$ es un* ***espacio afín*** *asociado a $\boldsymbol{E}$ si se tiene definida una aplicación:*

$$\psi: \mathcal{A}\times\mathcal{A} \to \boldsymbol{E}, \qquad (P,Q) \longmapsto \psi(P,Q) := \overrightarrow{PQ},$$

que satisface a los siguientes axiomas:

A_1') $\forall P\in\mathcal{A}, \psi_P: \mathcal{A}\to\boldsymbol{E}$, *definida por* $\psi_P(Q)=\psi(P,Q)$, *es biyectiva; es decir:* $\forall P\in\mathcal{A}, \forall\vec{\boldsymbol{v}}\in\boldsymbol{E},\ \dot{\exists} Q\in\mathcal{A} : \vec{\boldsymbol{v}}=\overrightarrow{PQ}$.

A_2') $\forall P,Q,R\in\mathcal{A} : \psi(P,Q)+\psi(Q,R)=\psi(P,R)$; *es decir:* $\overrightarrow{PQ}+\overrightarrow{QR}=\overrightarrow{PQ}$.

La estructura afín dada en $\mathcal{A}$ por esta definición la denotamos por $(\mathcal{A},\boldsymbol{E},\psi)$.

A partir de la estructura $(\mathcal{A},\boldsymbol{E},\phi)$, definimos $\psi:\mathcal{A}\times\mathcal{A}\to\boldsymbol{E}$ por $(P,Q)\mapsto\psi(P,Q)=\phi_P^{-1}(Q)$, que es único vector, $\overrightarrow{PQ}$, dado por el Axioma A_1 de la estructura $(\mathcal{A},\boldsymbol{E},\phi)$. Se verifican los Axiomas A_1' y A_2':

A_1': Como, para $\forall P\in\mathcal{A}, \psi_P=\phi_P^{-1}$, ψ_P es biyectiva.
A_2': Por el Axioma A_2 se tiene, $\forall P,Q,R\in\mathcal{A}$:

$$\psi(P,Q)+\psi(Q,R)=\phi_P^{-1}(Q)+\phi_Q^{-1}(R)\overset{(*)}{=}\phi_P^{-1}(R)=\psi(P,R).$$

Donde la igualdad central $(*)$, surge de aplicar ϕ_P a la suma de vectores anteriores al signo igual, y aplicar el Axioma A_2:

$$\phi_P\big(\phi_P^{-1}(Q)+\phi_Q^{-1}(R)\big)=\phi\big(P,\phi_P^{-1}(Q)+\phi_Q^{-1}(R)\big)=\phi\big(\phi(P,\phi_P^{-1}(Q)),\phi_Q^{-1}(R)\big)$$

$$=\phi\big(\phi_P(\phi_P^{-1}(Q)),\phi_Q^{-1}(R)\big)=\phi\big(Q,\phi_Q^{-1}(R)\big)=\phi_Q\big(\phi_Q^{-1}(R)\big)=R.$$

Así mismo, de la estructura afín definida en $\mathcal{A}$ por $(\mathcal{A},\boldsymbol{E},\psi)$ llegamos a la dada por la Definición 2.1, usando el Axioma A_1', para definir $\phi(P,\vec{\boldsymbol{v}})=\psi_P^{-1}(\vec{\boldsymbol{v}})$, $\forall P\in\mathcal{A},\vec{\boldsymbol{v}}\in\boldsymbol{E}$.

Una tercer enfoque para introducir al noción de espacio afín es el siguiente:

2.1" Definición.- *Dados un conjunto $\mathcal{A}$ no vacío y un espacio vectorial $\boldsymbol{E}$ sobre un cuerpo conmutativo K, diremos que $\mathcal{A}$ es un* ***espacio afín*** *asociado a $\boldsymbol{E}$ si, $\forall\vec{\boldsymbol{v}}\in\boldsymbol{E}$, se tiene definida una aplicación biyectiva:*

$$t_{\vec{\boldsymbol{v}}}:\mathcal{A}\to\mathcal{A}, \qquad P\longmapsto t_{\vec{\boldsymbol{v}}}(P),$$

verificándose:

A_1") $\forall P,Q\in\mathcal{A}, \dot{\exists}\vec{\boldsymbol{v}}\in\boldsymbol{E} : t_{\vec{\boldsymbol{v}}}(P)=Q$.
A_2") $\forall\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}\in\boldsymbol{E} : t_{\vec{\boldsymbol{u}}}\circ t_{\vec{\boldsymbol{v}}}=t_{\vec{\boldsymbol{u}}+\vec{\boldsymbol{v}}}$.

A cada elemento del conjunto

$$\mathfrak{T}=\{t_{\vec{\boldsymbol{u}}}:\mathcal{A}\to\mathcal{A}\ /\ \vec{\boldsymbol{u}}\in\boldsymbol{E}, \text{ verificándose los axiomas } A_1\text{" y } A_2\text{"}\}$$

les denominamos ***traslación*** de vector $\vec{\boldsymbol{u}}$. Y denotamos por $(\mathcal{A},\boldsymbol{E},\mathfrak{T})$ esta estructura afín en $\mathcal{A}$, que coincide con la estructura afín $(\mathcal{A},\boldsymbol{E},\phi)$, ya que $t_{\vec{\boldsymbol{v}}}(P)=\phi(P,\vec{\boldsymbol{v}})$, $\forall P\in\mathcal{A},\forall\vec{\boldsymbol{u}}\in\boldsymbol{E}$.

El conjunto $\mathfrak{T}$, de las traslaciones en el espacio afín $\mathcal{A}$, es un grupo conmutativo, respecto a la composición de aplicaciones ($t_{\vec{\boldsymbol{0}}} = I_{\mathcal{A}}, t_{\vec{\boldsymbol{v}}}^{-1} = t_{-\vec{\boldsymbol{v}}}$).

A partir de ahora, dadas las distintas caracterizaciones definidas de una estructura afín, usaremos indistintamente las notaciones introducidas en la Definición 2.1 y en las otras dos definiciones equivalentes dadas.

2.2 Definición.- *La **dimensión** de un espacio afín $\mathcal{A}$ es la dimensión del espacio vectorial $\boldsymbol{E}$ al cual está asociado.*

De los axiomas de espacio afín surgen las siguientes propiedades:

2.3 Proposición.- *Si P, Q, R y S son puntos de un espacio afín $\mathcal{A}$, se tienen las siguientes propiedades:*

1) $\overrightarrow{PP} = \vec{\boldsymbol{0}}$.
2) $\overrightarrow{PQ} = \vec{\boldsymbol{0}} \Rightarrow P = Q$.
3) $\overrightarrow{PQ} = -\overrightarrow{QP}$.
4) $\overrightarrow{PQ} = \overrightarrow{RS} \Leftrightarrow \overrightarrow{PR} = \overrightarrow{QS}$. *(Ley del paralelogramo)*
5) $\overrightarrow{PQ} + \overrightarrow{QR} + \overrightarrow{RP} = \vec{\boldsymbol{0}}$. *(Relación de Chasles)*

Demostración.- 1) Del Axioma A_2', $\overrightarrow{PP} + \overrightarrow{PP} = \overrightarrow{PP}$, luego $\overrightarrow{PP} = \vec{\boldsymbol{0}}$.

2) Como $\overrightarrow{PP} = \vec{\boldsymbol{0}}$ y si $\overrightarrow{PQ} = \vec{\boldsymbol{0}}$, se tiene que $\psi_P(P) = \psi_P(Q)$ y como (por el Axioma A_1') ψ es biyectiva, $P = Q$.

3) Del Axioma A_2', $\overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP} = \vec{\boldsymbol{0}}$, luego $\overrightarrow{PQ} = -\overrightarrow{QP}$.

4) Basta tener presente que $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$ y $\overrightarrow{QR} + \overrightarrow{RS} = \overrightarrow{QS}$, para que la equivalencia $\overrightarrow{PQ} = \overrightarrow{RS} \Leftrightarrow \overrightarrow{PR} = \overrightarrow{QS}$ se verifique.

5) Del Axioma A_2' se tiene que $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$, de donde $\overrightarrow{PQ} + \overrightarrow{QR} - \overrightarrow{PR} = \vec{\boldsymbol{0}}$ y de 3) de esta proposición, $\overrightarrow{PQ} + \overrightarrow{QR} + \overrightarrow{RP} = \vec{\boldsymbol{0}}$. ◾

Ejemplos de espacios afines

2.4 Ejemplo.- *Un espacio vectorial $\boldsymbol{E}$ es un espacio afín asociado a sí mismo (estructura afín canónica sobre $\boldsymbol{E}$).*

Basta tomar $(\mathcal{A} = \boldsymbol{E}, \boldsymbol{E}, +)$ siendo $+ \colon \mathcal{A} \times \boldsymbol{E} \to \mathcal{A}$ la propia suma de vectores en el espacio vectorial $\boldsymbol{E}$. Claramente se verifican los Axiomas A_1 y A_2 de espacio afín.

Si optamos por otra de las definiciones de estructura afín dada en la página 19, ponemos $\psi\colon \mathcal{A} \times \mathcal{A} \to \boldsymbol{E}$, $\psi(\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}) = \vec{\boldsymbol{v}} - \vec{\boldsymbol{u}}$.

2.5 Ejemplo.- *Sean K un cuerpo conmutativo y el espacio vectorial* (Ejemplo 1.3) *$\boldsymbol{E} = K^n$, en el conjunto $\mathcal{A} = K^n$ hay una **estructura afín canónica o estándar**.*

Para ello basta tomar cualquiera de las aplicaciones:

$$\begin{array}{ll} \phi\colon \mathcal{A}\times \boldsymbol{E}\to\mathcal{A} & ((\alpha_1,\dots,\alpha_n),(u_1,\dots,u_n))\mapsto(\alpha_1+u_1,\dots,\alpha_n+u_n)\\ \psi\colon \mathcal{A}\times\mathcal{A}\to\boldsymbol{E} & ((\alpha_1,\dots,\alpha_n),(\beta_1,\dots,\beta_n))\mapsto(\beta_1-\alpha_1,\dots,\beta_n-\alpha_n)\end{array}$$

2.6 Ejemplo.- *En el conjunto $\mathcal{P}_2[X]$ de los polinomios de grado 2 con coeficientes reales, se define una estructura de espacio afín $(\mathcal{P}_2[X],\mathbb{R}^3,\psi)$, definiendo*

$$\begin{array}{c}\psi\colon \mathcal{P}_2[X]\times\mathcal{P}_2[X]\to\mathbb{R}^3\\ (aX^2+bX+c,\bar{a}X^2+\bar{b}X+\bar{c})\longmapsto(\bar{a}-a,\bar{b}-b,\bar{c}-c).\end{array}$$

2.7 Ejemplo.- *Sean $\boldsymbol{F}$ un subespacio vectorial de un espacio vectorial $\boldsymbol{E}$ y $\vec{\boldsymbol{a}}\in\boldsymbol{E}$. El conjunto $\vec{\boldsymbol{a}}+\boldsymbol{F}=\{\vec{\boldsymbol{v}}\in\boldsymbol{E}\,/\,\vec{\boldsymbol{v}}=\vec{\boldsymbol{a}}+\vec{\boldsymbol{u}},\vec{\boldsymbol{u}}\in\boldsymbol{F}\}$ es un espacio afín asociado a $\boldsymbol{F}$.*

Una situación particular del este ejemplo se ha presentado en el Ejemplo 1.9 y la en Nota 1.10. Este ejemplo y el siguiente son situaciones particulares de subespacio afín o variedad lineal que se abordará en la Definición 2.13.

2.8 Ejemplo.- *Sean M es una matriz de orden $m\times n$ con coeficientes reales y B es una matriz columna de m números reales. El conjunto $\mathcal{A}$ de soluciones de la ecuación $MX=B$ es un espacio afín asociado al subespacio vectorial E de $\mathbb{R}^n$, conjunto de las soluciones del sistema homogéneo $MX=0$.*

En este ejemplo $MX=B$ son las ecuaciones cartesianas de $\mathcal{A}$, como variedad lineal (o subespacio afín) de $\mathbb{R}^n$, como veremos posteriormente (3-3).

Los dos ejemplos anteriores quedan englobados en el siguiente:

2.9 Ejemplo.- *Si $\boldsymbol{E}$ y $\boldsymbol{F}$ son espacios vectoriales, $f\in\mathcal{L}_K(\boldsymbol{E},\boldsymbol{F})$, $\vec{\boldsymbol{u}}_0\in\boldsymbol{E}$ y $\vec{\boldsymbol{v}}_0=f(\vec{\boldsymbol{u}}_0)$. El conjunto*

$$\mathcal{A}=\{\vec{\boldsymbol{u}}\in\boldsymbol{E}\,/\,f(\vec{\boldsymbol{u}})=\vec{\boldsymbol{v}}_0\}=\vec{\boldsymbol{u}}_0+\mathrm{Ker}(f)$$

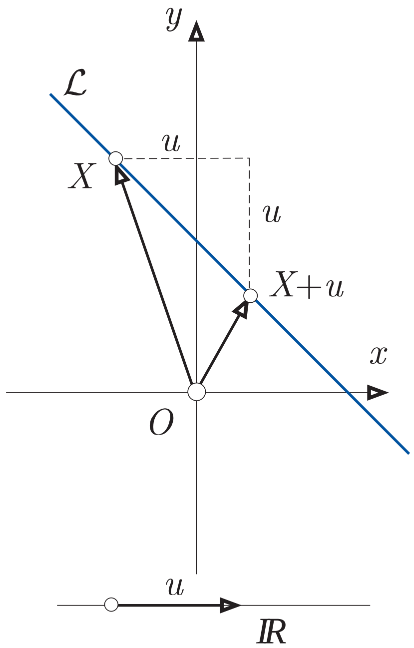
es un espacio afín asociado al $\mathrm{Ker}(f)$*.*

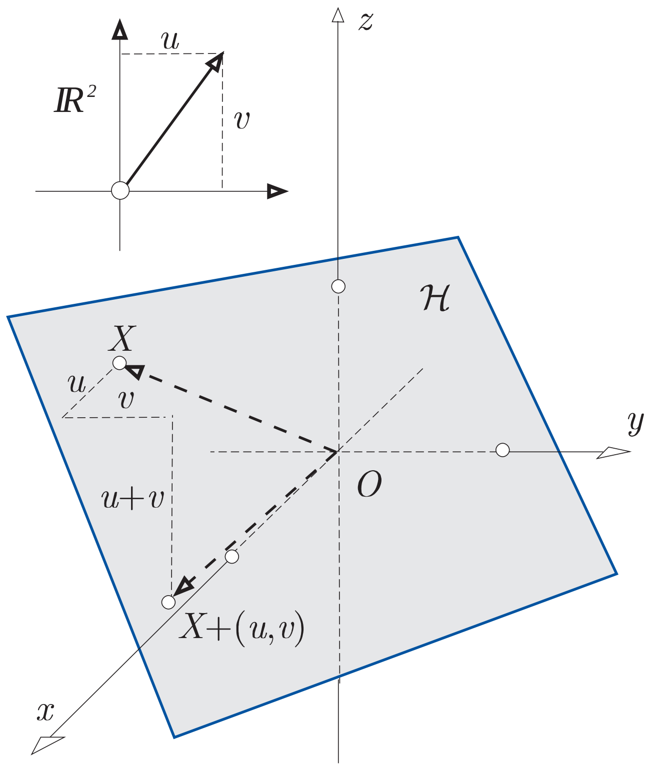
Vamos a dar tres ejemplos [4, pág 14] de espacios afines en los que no se pone de manifiesto (a menos de manera obvia) que su estructura es la inducida por la del espacio afín en el que están contenidos.

2.10 Ejemplo .- *Considérese el subconjunto $\mathfrak{L}\subset\mathbb{R}^2$ que consta de los puntos (x,y) que satisface la ecuación $x+y-1=0$; se trata de la recta de pendiente -1 y que pasa por los puntos $(1,0)$ y $(0,1)$.*

En $\mathfrak{L}$ definimos la operación externa $+\colon\mathfrak{L}\times\mathbb{R}\to\mathfrak{L}$, tal que para cada punto $(x,1-x)\in\mathfrak{L}$ y $\forall u\in\mathbb{R}$,

$$(x,1-x)+u=(x+u,1-x-u).$$





Es inmediato verificar que $(\mathcal{L}, I\!R, +)$ es una estructura afín en $\mathcal{L}$. Por ejemplo, dados dos puntos $P(a, 1-a)$ y $Q(b, 1-b)$ de $\mathcal{L}$, el único (vector) $u \in I\!R$ tal que $Q = P + u$ es $u = b - a$.

Este subconjunto $\mathcal{L} = \{(x, 1-x) \in I\!R^2 \ /x \in I\!R\} = (0, 1) + \boldsymbol{F}$ con $\boldsymbol{F}$ el subespacio vectorial $\{(x, -x) \in I\!R^2 \ /x \in I\!R\}$, tiene una estructura de espacio afín, como una situación particular del Ejemplo 2.7. También, $\mathcal{L}$ como subespacio afín de $I\!R^2$ (pág. 23).

2.11 Ejemplo.- *Sea $\mathcal{H} = \{(x, y, z) \in I\!R^3 \ /x + y + z = 1\}$, es decir, el plano que pasa por los puntos $(1, 0, 0), (0, 1, 0)$ y $(0, 0, 1)$.*
Damos en $\mathcal{H}$ una estructura afín, definiendo la operación
$+ \colon \mathcal{H} \times I\!R^2 \to \mathcal{H}$ *por:*
$(x, y, 1-x-y) + (u, v) = (x+u, y+v, 1-x-u-y-v).$

2.12 Ejemplo.- *Sea $\mathcal{P} = \{(x, y, z) \in I\!R^3 \ /x^2 + y^2 - z = 0\}$, paraboloide de revolución de eje OZ. A esta superficie se le puede dar estructura de espacio afín, con la operación externa $+ \colon \mathcal{P} \times I\!R^2 \to \mathcal{P}$, definida para todo punto $(x, y, x^2 + y^2) \in \mathcal{P}$ y todo $(u, v) \in I\!R^2$, por*

$$(x, y, x^2 + y^2) + (u, v) = (x + u, y + v, (x + u)^2 + (y + v)^2).$$

Estos tres últimos ejemplos de espacio afines ponen de manifiesto que no todos los espacios afines son obtenidos como subconjuntos de $I\!R^n$, con la estructura afín que la estructura canónica de $I\!R^n$ induce sobre ellos. Otros ejemplos, que no se ajustan a este tipo de espacios afines estándar, lo dan los espacios afines que se obtienen al quitar a un espacio proyectivo los puntos de uno cualquiera de sus hiperplanos [5, pág. 14].

2.2 Subespacios afines o variedades lineales afines

De nuevo, las rectas y planos en el espacio ordinario o bien el Ejemplo 2.7, nos sirven de modelo para dar la definición de subespacio afín. Debe ser el concepto que corresponde al de subespacio vectorial del espacio vectorial $\boldsymbol{E}$ al cual está asociado el espacio afín $(\mathcal{A}, \boldsymbol{E}, \phi)$ y, por tanto, ambos se deben relacionar mediante la aplicación biyectiva $\phi_P \colon \boldsymbol{E} \to \mathcal{A}$.

2.13 Definición.- *Sean $\mathcal{A}$ un espacio afín asociado a un espacio vectorial $\boldsymbol{E}$, $P \in \mathcal{A}$ y $\boldsymbol{F}$ un subespacio vectorial de $\boldsymbol{E}$, se llama **variedad lineal afín** (o simplemente **variedad lineal**), que pasa por P y de **subespacio director** $\boldsymbol{F}$, al conjunto*

$$\mathcal{F} = \{X \in \mathcal{A} \,/\, \overrightarrow{PX} \in \boldsymbol{F}\} = \{P + \vec{\boldsymbol{v}} \in \mathcal{A} \,/\, \vec{\boldsymbol{v}} \in \boldsymbol{F}\} := P + \boldsymbol{F}$$

2.14 Nota.- Obsérvese que $\mathcal{F} = \phi_P(\boldsymbol{F})$, siendo $(\mathcal{A}, \boldsymbol{E}, \phi)$ la estructura afín sobre $\mathcal{A}$; es decir, una variedad lineal es la imagen de un subespacio vectorial, mediante la aplicación biyectiva $\phi_P \colon \boldsymbol{E} \to \mathcal{A}$.

Dicho de otra forma, un subconjunto $\mathcal{F} \subset \mathcal{A}$ es una variedad lineal si $\exists P \in \mathcal{F}$ y un subespacio vectorial $\boldsymbol{F} \dot{\subset} \boldsymbol{E}$, tal que $\phi_P(\boldsymbol{F}) = \mathcal{F}$. En realidad el punto P que tomemos en $\mathcal{F}$ no es relevante: podemos coger cualquiera, como se verá en la Proposición 2.17.

En la Proposición 3.24 se da una caracterización de variedad lineal como el conjunto que contiene a la recta que pasa por cualquier par de sus puntos.

2.15 Definición.- *La **dimensión** de una variedad lineal es la del subespacio director. Las variedades lineales de dimensión 0 son los puntos ($P + \{\vec{\boldsymbol{0}}\} = \{P\}$). Las variedades lineales de dimensión 1 se llaman **rectas**; puntos contenidos en una misma recta, se dicen que están **alineados**. Las variedades lineales de dimensión 2, se llaman **planos**; puntos pertenecientes a un mismo plano se dicen que son **coplanarios**. Si $\dim \boldsymbol{E} = n$, las variedades lineales de dimensión $n-1$ se llaman **hiperplanos**.*

A las variedades lineales se les conoce también como ***subespacios afines***, la razón es que una variedad lineal $\mathcal{F} \dot{\subset} \mathcal{A}$ es un espacio afín, con la estructura inducida por el espacio afín $(\mathcal{A}, \boldsymbol{E}, \psi)$:

$$\psi_{|\mathcal{F}\times\mathcal{F}} \colon \mathcal{F} \times \mathcal{F} \to \boldsymbol{F}$$

La variedad lineal $\mathcal{F} = P + \boldsymbol{F}$ de un espacio afín $\mathcal{A}$ no depende del punto P elegido: podemos tomar otro punto Q de ella para describirla; es decir:

2.16 Proposición.- *Sean $(\mathcal{A}, \boldsymbol{E}, +)$ un espacio afín, $P \in \mathcal{A}$, $\boldsymbol{F}$ es un subespacio vectorial de $\boldsymbol{E}$ y $Q \in \mathcal{F} = P + \boldsymbol{F}$, entonces $\mathcal{F} = Q + \boldsymbol{F}$.*

Demostración.- Como $Q \in P + \boldsymbol{F}$, se tiene $\overrightarrow{PQ} \in \boldsymbol{F}$ y $-\overrightarrow{PQ} = \overrightarrow{QP} \in \boldsymbol{F}$.
Para probar que $P + \boldsymbol{F} = Q + \boldsymbol{F}$, establezcamos la doble inclusión:

$$X \in P + \boldsymbol{F} \Rightarrow \overrightarrow{PX} \in \boldsymbol{F} \Rightarrow \overrightarrow{QP} + \overrightarrow{PX} = \overrightarrow{QX} \in \boldsymbol{F} \Rightarrow X \in Q + \boldsymbol{F}.$$
$$X \in Q + \boldsymbol{F} \Rightarrow \overrightarrow{QX} \in \boldsymbol{F} \Rightarrow \overrightarrow{PQ} + \overrightarrow{QX} = \overrightarrow{PX} \in \boldsymbol{F} \Rightarrow X \in P + \boldsymbol{F}.$$

Tenemos, entonces, la caracterización de una variedad lineal dada por la proposición siguiente:

2.17 Proposición.- *Si $\mathcal{F}$ es una variedad lineal de un espacio afín $(\mathcal{A}, \boldsymbol{E}, +)$ de subespacio director $\boldsymbol{F} \dot{\subset} \boldsymbol{E}$, entonces $\boldsymbol{F} = \{\overrightarrow{PQ} \in \boldsymbol{E} \,/ P, Q \in \mathcal{F}\}$.*

Demostración.- El que $\boldsymbol{F}$ sea el subespacio director de $\mathcal{F}$ significa que existe un punto $P_0 \in \mathcal{F}$ tal que $\mathcal{F} = P_0 + \boldsymbol{F}$.

Si $P, Q \in \mathcal{F} = P_0 + \boldsymbol{F} \Rightarrow \overrightarrow{P_0P} \in \boldsymbol{F}, \overrightarrow{P_0Q} \in \boldsymbol{F} \Rightarrow \overrightarrow{PP_0} + \overrightarrow{P_0Q} = \overrightarrow{PQ} \in \boldsymbol{F}$.

Para la otra inclusión, si $\vec{\boldsymbol{u}} \in \boldsymbol{F}$, el punto $P = P_0 + \vec{\boldsymbol{u}} \in \mathcal{F} \Rightarrow \overrightarrow{P_0P} = \vec{\boldsymbol{u}} \Rightarrow \vec{\boldsymbol{u}} \in \{\overrightarrow{PQ} \in \boldsymbol{E} \,/ P, Q \in \mathcal{F}\}$.

Esta proposición nos da otra definición alternativa de variedad lineal (posteriormente, daremos otra definición equivalente en términos de coordenadas baricéntricas (Proposición 3.24), de la que se ha mencionado un caso particular en la Nota 1.10):

2.18 Definición.- *Un subconjunto $\mathcal{F}$ de un espacio afín $\mathcal{A}$ asociado a un espacio vectorial $\boldsymbol{E}$ es una variedad lineal si $\{\overrightarrow{PQ} \in \boldsymbol{E} \,/ P, Q \in \mathcal{F}\}$ es un subespacio vectorial de $\boldsymbol{E}$.*

Para comprobar que un subconjunto $\mathcal{F}$ de un espacio afín $\mathcal{A}$ asociado a un espacio vectorial $\boldsymbol{E}$ es una variedad lineal basta tomar un punto $P_0 \in \mathcal{F}$ y establecer que $\{\overrightarrow{P_0P} \in \boldsymbol{E} \,/ P \in \mathcal{F}\}$ es un subespacio vectorial de $\boldsymbol{E}$. En efecto, $\{\overrightarrow{PQ} \in \boldsymbol{E} \,/ P, Q \in \mathcal{F}\} = \{\overrightarrow{P_0P} \in \boldsymbol{E} \,/ P \in \mathcal{F}\}$, pues $\overrightarrow{PQ} = \overrightarrow{PP_0} + \overrightarrow{P_0Q}$, por el Axioma A_2.

2.19 Ejemplo.- *La recta $\mathcal{L} = \{(x, y) \in \mathbb{R}^2 \,/ x + y = 1\}$ del plano afín real, $(\mathbb{R}^2, \mathbb{R}^2, +)$, es una variedad lineal $\mathcal{L} = P + \boldsymbol{L}$, tomando el punto P, por ejemplo, como $P = (1, 0)$ y $\boldsymbol{L} = \{(x, -x) \in \mathbb{R}^2 \,/ x \in \mathbb{R}\}$ subespacio vectorial de $\mathbb{R}^2$.*

De hecho cualquier recta del plano afín real es así mismo un espacio afín, con la estructura afín $(\mathcal{L}, \boldsymbol{L}, +)$ que el plano afín real induce sobre ella.

2.20 Nota.- El que un subconjunto A no vacío de un espacio afín $\mathcal{A}$ sea una variedad lineal se puede caracterizar (ver pág. 48) exigiendo que contenga al baricentro de cualquier subconjunto de él.
Como consecuencia resulta que un subconjunto (con más de un punto) de un espacio afín $\mathcal{A}$ que contiene a cada recta que pasa por dos cualesquiera de sus puntos es una variedad lineal.

2.3 Intersección y suma de variedades lineales. Paralelismo

Vamos a estudiar la posición relativa de dos variedades lineales, analizando primero cuando tienen puntos comunes.

2.21 Proposición.- *Sean $P, Q \in \mathcal{A}$ dos puntos de un espacio afín asociado a un espacio vectorial $\boldsymbol{E}$, $\boldsymbol{F} \dot{\subset} \boldsymbol{E}$ y $\boldsymbol{G} \dot{\subset} \boldsymbol{E}$ subespacios vectoriales de $\boldsymbol{E} \Rightarrow$*

$$\left[(P+\boldsymbol{F}) \cap (Q+\boldsymbol{G}) \neq \varnothing \Leftrightarrow \overrightarrow{PQ} \in \boldsymbol{F}+\boldsymbol{G}\right].$$

Demostración.-
Si $(P+\boldsymbol{F}) \cap (Q+\boldsymbol{G}) \neq \varnothing \Rightarrow \exists R \in \mathcal{A} : R \in P+\boldsymbol{F}, R \in Q+\boldsymbol{G} \Rightarrow$
$\overrightarrow{PR} \in \boldsymbol{F}, \overrightarrow{QR} \in \boldsymbol{G} \Rightarrow \overrightarrow{PQ} = \overrightarrow{PR} + \overrightarrow{RQ} = \overrightarrow{PR} - \overrightarrow{QR} \in \boldsymbol{F}+\boldsymbol{G}$.
Si $\overrightarrow{PQ} \in \boldsymbol{F}+\boldsymbol{G} \Rightarrow \exists \vec{\boldsymbol{u}} \in \boldsymbol{F}, \exists \vec{\boldsymbol{v}} \in \boldsymbol{G} : \overrightarrow{PQ} = \vec{\boldsymbol{u}} + \vec{\boldsymbol{v}} \Rightarrow$
$P + \vec{\boldsymbol{u}} \in P+\boldsymbol{F}, \quad P+\vec{\boldsymbol{u}} = P + (\overrightarrow{PQ} - \vec{\boldsymbol{v}}) = Q + (-\vec{\boldsymbol{v}}) \in Q+\boldsymbol{G} \Rightarrow$
$P + \vec{\boldsymbol{u}} \in (P+\boldsymbol{F}) \cap (Q+\boldsymbol{G}) \neq \varnothing$. ◼

2.22 Ejemplo.- *En el espacio afín $I\!R^4$, con la estructura canónica, no tienen puntos comunes los subconjuntos:*

$$\mathcal{F} = \{(x,1,x+t,t) \in I\!R^4 \;/ x \in I\!R, t \in I\!R\},$$
$$\mathcal{G} = \{(\bar{x},\bar{y},\bar{x}-1,1) \in I\!R^4 \;/ \bar{x} \in I\!R, \bar{y} \in I\!R\}.$$

Ambos son variedades lineales (planos) ya que $\mathcal{F} = P_0 + \boldsymbol{F}$ y $\mathcal{G} = Q_0 + \boldsymbol{G}$, siendo $P_0 = (0,1,0,0)$, $Q_0 = (0,0,-1,1)$, y $\boldsymbol{F}$ y $\boldsymbol{G}$ son los subespacio vectoriales de $I\!R^4$:
$\boldsymbol{F} = \{(x,0,x+t,t) \in I\!R^4 \;/ x \in I\!R, t \in I\!R\}$, $\boldsymbol{G} = \{(\bar{x},\bar{y},\bar{x},0) \in I\!R^4 \;/ x \in I\!R, t \in I\!R\}$.

$\mathcal{F} \cap \mathcal{G} = \varnothing$, pues el vector $\overrightarrow{P_0Q_0} = (0,-1,-1,1) \notin \boldsymbol{F}+\boldsymbol{G}$. Ya que no se puede poner como una suma $\vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}$, con $\vec{\boldsymbol{u}} = (x,0,x+t,t)$ y $\vec{\boldsymbol{v}} = (\bar{x},\bar{y},\bar{x},0)$; pues, el sistema de ecuaciones siguiente no tiene solución:

$$x+\bar{x} = 0, \quad \bar{y} = -1, \quad x+t+\bar{x} = -1, \quad t = 1.$$

Claramente, llegamos también a que $\mathcal{F} \cap \mathcal{G} = \varnothing$, comprobando que el sistema de ecuaciones, formado por las ecuaciones de ambas variedades lineales [(2)] no tiene solución:

$$z = x+y, \quad y = 1; \qquad x - z = 1, \quad t = 1.$$

Si dos variedades lineales son disjuntas, la unión de sus subespacios directores no generan el espacio vectorial asociado al espacio afín donde están contenidas; concretando:

2.23 Proposición.- *Sean $(\mathcal{A}, \boldsymbol{E})$ un espacio afín, $\mathcal{F}$ y $\mathcal{G}$ dos variedades lineales no vacías con subespacios directores $\boldsymbol{F}$ y $\boldsymbol{G}$, respectivamente, entonces:*

$$(\mathcal{F} \cap \mathcal{G} = \varnothing \Rightarrow \boldsymbol{F}+\boldsymbol{G} \subsetneqq \boldsymbol{E}) \Leftrightarrow (\boldsymbol{F}+\boldsymbol{G} = \boldsymbol{E} \Rightarrow \mathcal{F} \cap \mathcal{G} \neq \varnothing).$$

(2) Daremos un tratamiento general sobre las ecuaciones de las variedades lineales en el § 3.1.

Demostración.- Sean $P \in \mathcal{F}$ y $Q \in \mathcal{G}$, entonces $\overrightarrow{PQ} \in \boldsymbol{E}$ y si $\boldsymbol{E} = \boldsymbol{F} + \boldsymbol{G}$, existen $\vec{\boldsymbol{u}} \in \boldsymbol{F}$ y $\vec{\boldsymbol{v}} \in \boldsymbol{G}$ tales que $\overrightarrow{PQ} = \vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}$.

$Q + (-\vec{\boldsymbol{v}}) \in \mathcal{G}$ y $Q + (-\vec{\boldsymbol{v}}) = ((P + \vec{\boldsymbol{u}} + \vec{\boldsymbol{v}})) + (-\vec{\boldsymbol{v}}) = P + \vec{\boldsymbol{u}} \in \mathcal{F}$; luego, $\mathcal{F} \cap \mathcal{G} \neq \varnothing$ ◪

Veamos ahora quién es el subespacio director de la intersección de variedades lineales.

2.24 Proposición.- *Sean $\mathcal{F}_i = P_i + \boldsymbol{F}_i$, $(i \in I)$ variedades lineales, con intersección no vacía, y $P \in \bigcap_{i\in I} \mathcal{F}_i$, entonces $\bigcap_{i\in I} \mathcal{F}_i$ es una variedad que pasa por P y de subespacio director $\bigcap_{i\in I} \boldsymbol{F}_i$:*

$$\bigcap_{i\in I} \mathcal{F}_i = P + \bigcap_{i\in I} \boldsymbol{F}_i.$$

Demostración.-

$$Q \in P + \bigcap_{i\in I} \boldsymbol{F}_i \Leftrightarrow \overrightarrow{PQ} \in \bigcap_{i\in I} \boldsymbol{F}_i \Leftrightarrow (\forall i \in I : \overrightarrow{PQ} \in \boldsymbol{F}_i) \Leftrightarrow$$
$$\Leftrightarrow (\forall i \in I : Q = P + \overrightarrow{PQ} \in \mathcal{F}_i) \Leftrightarrow Q \in \bigcap_{i\in I} \mathcal{F}_i.$$ ◪

2.25 Ejemplo.- *Dos variedades lineales $P + \boldsymbol{F}$ y $Q + \boldsymbol{G}$ de un espacio afín $\mathcal{A}$, asociado a un espacio vectorial $\boldsymbol{E}$, de subespacio directores suplementarios ($\boldsymbol{F} + \boldsymbol{G} = \boldsymbol{E}$ y $\boldsymbol{F} \cap \boldsymbol{G} = \{\vec{\boldsymbol{0}}\}$, i.e. $\boldsymbol{E} = \boldsymbol{F} \oplus \boldsymbol{G}$) tienen un único punto común.*

$\overrightarrow{PQ} \in \boldsymbol{E} = \boldsymbol{F} + \boldsymbol{G} \overset{\text{Proposición 2.21}}{\Longrightarrow} (P + \boldsymbol{F}) \cap (Q + \boldsymbol{G}) \neq \varnothing \Rightarrow$

$\exists R \in (P + \boldsymbol{F}) \cap (Q + \boldsymbol{G}) \overset{\text{Proposición 2.24}}{\Longrightarrow} (P + \boldsymbol{F}) \cap (Q + \boldsymbol{G}) = R + (\boldsymbol{F} \cap \boldsymbol{G}) = R + \{\vec{\boldsymbol{0}}\} = \{R\}$.

Esto ocurre, en particular, en el caso de una recta y un hiperplano, que no contenga a la recta, en cualquier espacio afín.

Variedad lineal generada por un subconjunto de puntos

Tratamos ahora de encontrar la variedad lineal minimal que contiene a un subconjunto de puntos de un espacio afín, para ello usaremos el concepto de subespacio vectorial generado (Proposición 1.12) por un subconjunto de un espacio vectorial.

Sean $\mathcal{C}$ un subconjunto no vacío de un espacio afín $\mathcal{A}$, asociado a un espacio vectorial $\boldsymbol{E}$, y $\mathfrak{F}(\mathcal{C}) = \{\mathcal{F} \dot{\subset} \mathcal{A} \ / \mathcal{F}$ es una variedad lineal y $\mathcal{C} \subset \mathcal{F}\}$.

2.26 Proposición.- *Sean $P_0 \in \mathcal{C} \subset \mathcal{A}$ y $\boldsymbol{F}$ un subespacio vectorial de $\boldsymbol{E}$ $\Rightarrow \left[\mathcal{F} = P_0 + \boldsymbol{F} \in \mathfrak{F}(\mathcal{C}) \Leftrightarrow \forall P, Q \in \mathcal{C} : \overrightarrow{PQ} \in \boldsymbol{F}\right]$.*

Demostración.- Si $\mathcal{F} = P_0 + \boldsymbol{F} \in \mathfrak{F}(\mathcal{C})$ y $P, Q \in \mathcal{C} \Rightarrow P, Q \in \mathcal{F} \Rightarrow \overrightarrow{PQ} \in \boldsymbol{F}$.

Sean $\mathcal{F} = P_0 + \boldsymbol{F}$ una variedad lineal y se verifica que $\forall P, Q \in \mathcal{C}, \overrightarrow{PQ} \in \boldsymbol{F}$, demostremos que $\mathcal{F} \in \mathfrak{F}(\mathcal{C})$:

$$\text{Si } P \in \mathcal{C} \Rightarrow \overrightarrow{P_0P} \in \boldsymbol{F} \Rightarrow P \in P_0 + \boldsymbol{F} = \mathcal{F} \Rightarrow \mathcal{F} \in \mathfrak{F}(\mathcal{C}).$$ ◪

Como consecuencia de esta proposición podemos afirmar:

"Si $\mathcal{C}$ es un subconjunto no vacío de un espacio afín $\mathcal{A}$ asociado a un espacio vectorial $\mathbf{E}$*,* $\mathsf{C} = \{\overrightarrow{PQ} \in \mathbf{E} \,/ P, Q \in \mathcal{C}\}$*,* $\widetilde{\mathsf{C}}$ *el subespacio vectorial de* $\mathbf{E}$ *generado por los vectores de* C *(menor subespacio vectorial que contiene a* C*) y* $P_0 \in \mathcal{C}$*, entonces la variedad lineal* $\mathcal{F} = P_0 + \widetilde{\mathsf{C}}$ *es la menor que contiene a* $\mathcal{C}$*".*

En efecto, cualquier otra variedad lineal $\mathcal{F}$' que contenga a $\mathcal{C}$ sería de la forma $\mathcal{F}' = P_0 + \boldsymbol{F}'$, con $\boldsymbol{F}'$ subespacio vectorial de $\boldsymbol{E}$ que contiene a $\widetilde{C}$; luego, $\mathcal{F} \subset \mathcal{F}'$.

2.27 Definición.- *Si* $P_0 \in \mathcal{C} \subset \mathcal{A}$ *y* $\mathsf{C} = \{\overrightarrow{PQ} \in \mathbf{E} \,/ P, Q \in \mathcal{C}\}$*,* $\widetilde{\mathsf{C}}$ *el subespacio vectorial de* $\boldsymbol{E}$ *generado por los vectores de* C*, a la variedad lineal* $P_0 + \widetilde{\mathsf{C}}$ *se le denomina* ***variedad lineal generada por*** $\mathcal{C}$ *y la denotamos por* $\widetilde{\mathcal{C}} = P_0 + \widetilde{\mathsf{C}}$*.*

2.28 Ejemplo.- *Por dos puntos distintos de un espacio afín pasa una y solo una recta.*

Dados dos puntos P y Q de un espacio afín $\mathcal{A}$, la variedad lineal generada por $\{P, Q\}$ (menor variedad lineal que lo contiene) tiene como subespacio director el generado por $\{\overrightarrow{PQ}\}$, que es de dimensión 1. Designaremos por PQ a la recta que pasa por dos puntos P y Q.

Este es el primer Postulado de Euclides: *Dados dos puntos se puede trazar una y solo una recta que los contiene.*

Suma de variedades lineales

Si bien la intersección de variedades lineales es una variedad lineal, no ocurre lo mismo con la unión de variedades lineales.

Por ejemplo, sean $\mathcal{L}_i = P + \boldsymbol{L}_i$ $(i = 1, 2)$ dos rectas que pasan por $P \in \mathcal{A}$ y tienen subespacio directores unidimensionales distintos. Si $\mathcal{L}_1 \cup \mathcal{L}_2$ es una variedad lineal existe un subespacio vectorial $\boldsymbol{F}$ y $\mathcal{L}_1 \cup \mathcal{L}_2 = P + \boldsymbol{F}$. Si $P_1 \in \mathcal{L}_1 \subset \mathcal{L}_1 \cup \mathcal{L}_2$ $(P \neq P_1)$ y $P_2 \in \mathcal{L}_2 \subset \mathcal{L}_1 \cup \mathcal{L}_2$ $(P \neq P_2)$, $\overrightarrow{P_1P_2} \in \boldsymbol{F}$. Luego, el punto $Q = P + \overrightarrow{P_1P_2} \in \mathcal{L}_1 \cup \mathcal{L}_2$. Sin embargo, $Q = P + \overrightarrow{P_1P_2} = P + (\overrightarrow{P_1P} + \overrightarrow{PP_2})$ no pertenece ni a $\mathcal{L}_1$ ni a $\mathcal{L}_2$; ya que, $\overrightarrow{P_1P} + \overrightarrow{PP_2} \notin \boldsymbol{L}_1$ y $\overrightarrow{P_1P} + \overrightarrow{PP_2} \notin \boldsymbol{L}_2$, por ser $\overrightarrow{P_1P}$, $\overrightarrow{PP_2}$ vectores linealmente independientes y $\boldsymbol{L}_1$ y $\boldsymbol{L}_2$ son unidimensionales.

Sin embargo, si $\mathsf{C} = \{\overrightarrow{PQ} \in \boldsymbol{E} \,/ P, Q \in \mathcal{L}_1 \cup \mathcal{L}_2\}$ entonces $\mathcal{C} = P + \widetilde{\mathsf{C}}$ si es una variedad lineal y es la menor variedad lineal que contiene a $\mathcal{L}_1 \cup \mathcal{L}_2$.

Definimos la suma de variedades lineales como la menor variedad lineal que contiene a la unión de ellas; más precisamente:

2.29 Definición.- *Sean* $\mathcal{F}_1$ *y* $\mathcal{F}_2$ *variedades lineales de un espacio afín* $\mathcal{A}$*, se llama* ***suma de las variedades lineales*** $\mathcal{F}_1$ *y* $\mathcal{F}_2$ *a la variedad lineal generada por el conjunto unión de todas ellas:*

$$\mathcal{F}_1 + \mathcal{F}_2 = \widetilde{\mathcal{F}_1 \cup \mathcal{F}_i}\,.$$

Si las variedades lineales $\mathcal{F}_i$ tienen subespacios directores $\boldsymbol{F}_i$ $(i=1,2)$, el subespacio director de $\mathcal{F}_1+\mathcal{F}_2$ no es $\boldsymbol{F}_1+\boldsymbol{F}_2$, a menos que $\mathcal{F}_1\cap\mathcal{F}_2\neq\varnothing$:

2.30 Proposición.- *Sean $\mathcal{F}_1,\mathcal{F}_2$ variedades lineales de un espacio afín $\mathcal{A}$ de intersección NO vacía, con subespacios directores $\boldsymbol{F}_1,\boldsymbol{F}_2$ y $P_0\in\mathcal{F}_1\cap\mathcal{F}_2$, entonces:*

$$\mathcal{F}_1+\mathcal{F}_2=P_0+(\boldsymbol{F}_1+\boldsymbol{F}_2).$$

Demostración.- La variedad lineal $P_0+(\boldsymbol{F}_1+\boldsymbol{F}_2)$ contiene a $\mathcal{F}_1\cup\mathcal{F}_2$:

$$P\in\mathcal{F}_1\cup\mathcal{F}_2\Rightarrow\left\{\begin{array}{c} P\in\mathcal{F}_1=P_0+\boldsymbol{F}_1\subset P_0+(\boldsymbol{F}_1+\boldsymbol{F}_2)\\ \text{ó}\\ P\in\mathcal{F}_2=P_0+\boldsymbol{F}_2\subset P_0+(\boldsymbol{F}_1+\boldsymbol{F}_2)\end{array}\right\}\Rightarrow P\in P_0+(\boldsymbol{F}_1+\boldsymbol{F}_2).$$

Ahora, como $\mathcal{F}_1+\mathcal{F}_2$ es la menor variedad lineal que contiene a $\mathcal{F}_1\cup\mathcal{F}_2$, se tiene una inclusión: $\mathcal{F}_1+\mathcal{F}_2\subset P_0+(\boldsymbol{F}_1+\boldsymbol{F}_2)$.

Para la otra inclusión, sea $P\in P_0+(\boldsymbol{F}_1+\boldsymbol{F}_2)$, entonces $P=P_0+(\vec{\boldsymbol{v}}_1+\vec{\boldsymbol{v}}_2)$, con $\vec{\boldsymbol{v}}_1\in\boldsymbol{F}_1$ y $\vec{\boldsymbol{v}}_2\in\boldsymbol{F}_2$. Se tiene que, para $i=1,2$:

$$P_i=P_0+\vec{\boldsymbol{v}}_i\in\mathcal{F}_i\subset\mathcal{F}_1\cup\mathcal{F}_2\subset\widetilde{\mathcal{F}_1\cup\mathcal{F}_i}=\mathcal{F}_1+\mathcal{F}_2.$$

Luego, los vectores $\vec{\boldsymbol{v}}_i=\overrightarrow{P_0P_i}$ pertenecen al subespacio director de $\mathcal{F}_1+\mathcal{F}_2$; y también $\vec{\boldsymbol{v}}_1+\vec{\boldsymbol{v}}_2$. Así, definitivamente, $P=P_0+(\vec{\boldsymbol{v}}_1+\vec{\boldsymbol{v}}_2)\in\mathcal{F}_1+\mathcal{F}_2$. ◪

2.31 Proposición.- *Sean $\mathcal{F}_1=P+\boldsymbol{F}_1$ y $\mathcal{F}_2=Q+\boldsymbol{F}_2$ variedades lineales de un espacio afín $\mathcal{A}$ de intersección vacía, entonces:*

$$\mathcal{F}_1+\mathcal{F}_2=P+(\boldsymbol{F}_1+\boldsymbol{F}_2+\widetilde{\{\overrightarrow{PQ}\}}).$$

Demostración.- Como $P,Q\in P+(\boldsymbol{F}_1+\boldsymbol{F}_2+\widetilde{\{\overrightarrow{PQ}\}})$, esta variedad lineal contiene a $\mathcal{F}_1=P+\boldsymbol{F}_1$ y a $\mathcal{F}_2=Q+\boldsymbol{F}_2$ y por tanto, a la unión de ambas. Y al ser $\mathcal{F}_1+\mathcal{F}_2$ la menor variedad lineal que contiene a $\mathcal{F}_1\cup\mathcal{F}_2$:

$$\mathcal{F}_1+\mathcal{F}_2\subset P+(\boldsymbol{F}_1+\boldsymbol{F}_2+\widetilde{\{\overrightarrow{PQ}\}}).$$

Para obtener la otra inclusión, si $\boldsymbol{F}$ es el subespacio director de la variedad lineal suma de $\mathcal{F}_1$ y $\mathcal{F}_2$ $(\mathcal{F}_1+\mathcal{F}_2=P+\boldsymbol{F})$ y como $P,Q\in P+\boldsymbol{F}\Rightarrow\overrightarrow{PQ}\in\boldsymbol{F}$.

$X\in P+(\boldsymbol{F}_1+\boldsymbol{F}_2+\widetilde{\{\overrightarrow{PQ}\}})\Rightarrow\overrightarrow{PX}\in\boldsymbol{F}_1+\boldsymbol{F}_2+\widetilde{\{\overrightarrow{PQ}\}}\Rightarrow$
$\overrightarrow{PX}=\vec{\boldsymbol{v}}_1+\vec{\boldsymbol{v}}_2+\lambda\overrightarrow{PQ}\quad(\vec{\boldsymbol{v}}_1\in\boldsymbol{F}_1,\vec{\boldsymbol{v}}_2\in\boldsymbol{F}_2,\lambda\in K)$
$P_1=P+\vec{\boldsymbol{v}}_1\in P+\boldsymbol{F}_1\subset P+\boldsymbol{F}\Rightarrow\vec{\boldsymbol{v}}_1=\overrightarrow{PP_1}\in\boldsymbol{F}$.
$P_2=Q+\vec{\boldsymbol{v}}_2\in Q+\boldsymbol{F}_2\subset P+\boldsymbol{F}\Rightarrow\overrightarrow{PP_2}\in\boldsymbol{F}\Rightarrow\vec{\boldsymbol{v}}_2=\overrightarrow{QP_2}=\overrightarrow{PP_2}-\overrightarrow{PQ}\in\boldsymbol{F}$.
Luego, $\overrightarrow{PX}\in\boldsymbol{F}\Rightarrow X\in P+\boldsymbol{F}$. Concluimos que:

$$P+(\boldsymbol{F}_1+\boldsymbol{F}_2+\widetilde{\{\overrightarrow{PQ}\}})\subset P+\boldsymbol{F}.$$ ◪

Las dos proposiciones anteriores, junto con la fórmula de la dimensión de la suma de subespacios vectoriales (Proposición 1.27), nos permiten obtener la dimensión de la variedad lineal suma de otras dos.

2.32 Proposición.- *Sean $\mathcal{F}_1 = P + F_1$ y $\mathcal{F}_2 = Q + F_2$ variedades lineales de un espacio afín $\mathcal{A}$, entonces:*
Si $\mathcal{F}_1 \cap \mathcal{F}_2 \neq \varnothing$, $\dim(\mathcal{F}_1 + \mathcal{F}_2) = \dim \mathcal{F}_1 + \dim \mathcal{F}_2 - \dim(\mathcal{F}_1 \cap \mathcal{F}_2)$.
Si $\mathcal{F}_1 \cap \mathcal{F}_2 = \varnothing$, $\dim(\mathcal{F}_1 + \mathcal{F}_2) = \dim \mathcal{F}_1 + \dim \mathcal{F}_2 - \dim(F_1 \cap F_2) + 1$.

Demostración.- La dimensión de una variedad lineal es la de su subespacio director; así, si $\mathcal{F}_1 \cap \mathcal{F}_2 \neq \varnothing$, por Proposición 1.27, Proposición 2.24 y Proposición 2.30:

$$\dim(\mathcal{F}_1 + \mathcal{F}_2) = \dim \mathcal{F}_1 + \dim \mathcal{F}_2 - \dim(\mathcal{F}_1 \cap \mathcal{F}_2).$$

Si $\mathcal{F}_1 \cap \mathcal{F}_2 = \varnothing$, por la Proposición 2.21, se tiene que $\overrightarrow{PQ} \notin F_1 + F_2$; y por la Proposición 2.31, $\dim(\mathcal{F}_1 + \mathcal{F}_2) = \dim(F_1 + F_2) + 1$; por tanto:

$$\dim(\mathcal{F}_1 + \mathcal{F}_2) = \dim \mathcal{F}_1 + \dim \mathcal{F}_2 - \dim(F_1 \cap F_2) + 1.$$ ◪

Variedades lineales paralelas

2.33 Definición.- *Se dice que dos variedades lineales, $\mathcal{F} = P+F$ y $\mathcal{G} = Q+G$, son **paralelas** (se designa $\mathcal{F} \| \mathcal{G}$) si $F \subseteq G$ ó $G \subseteq F$.*

2.34 Definición.- *Se dice que dos variedades lineales, $\mathcal{F} = P+F$ y $\mathcal{G} = Q+G$, se **cruzan** si $\mathcal{F} \cap \mathcal{G} = \varnothing$ y $F \cap G = \{\vec{0}\}$.*

2.35 Ejemplo.- *En espacio afín ordinario:*
— Dos rectas son paralelas si sus espacios directores coinciden.
— Dos planos son paralelos si sus espacios directores coinciden.
— Una recta y un plano son paralelos si el subespacio director de la recta está contenido en el subespacio director del plano.

2.36 Ejemplo.- *Dos rectas distintas en un plano afín son paralelas o tienen un único punto común.*

Sean $\mathcal{L}_1 = P+L_1$ y $\mathcal{L}_2 = Q+L_2$ dos rectas (variedades lineales de dimensión 1) de un plano afín $\mathcal{A}$ ($\dim \mathcal{A} = 2$) asociado a un espacio vectorial E. Si $\mathcal{L}_1 \nparallel \mathcal{L}_2 \Rightarrow L_1 \cap L_2 = \{\vec{0}\}$ y $L_1 + L_2 = E$. Luego L_1 y L_2 son suplementarios y del Ejemplo 2.25, se sigue que tienen un sólo punto común.

También podemos verificar este hecho, usando las fórmulas 2.32 de la dimensión de la suma de variedades lineales. Supongamos que las dos rectas no se cortan, entonces $\dim(\mathcal{L}_1 + \mathcal{L}_2) = \dim \mathcal{L}_1 + \dim \mathcal{L}_2 - \dim(L_1 \cap L_2) + 1 = 3 - \dim(L_1 \cap L_2)$. Por tanto, $\dim(L_1 \cap L_2) \geq 1$, pero como $\dim L_1 = \dim L_2 = 1$, ocurre que $\dim(L_1 \cap L_2) = 1$, es decir, $L_1 = L_2 \Rightarrow \mathcal{L}_1 \| \mathcal{L}_2$.

2.37 Ejemplo.- *En un espacio afín tridimensional un recta y un plano no paralelos se cortan en un único punto.*

2.38 Ejemplo.- *Dos variedades lineales paralelas o no se cortan o una está contenida en la otra.*

Sean $\mathcal{F} = P + \boldsymbol{F}$ y $\mathcal{G} = Q + \boldsymbol{G}$ dos variedades lineales de un espacio afín $\mathcal{A}$ tales que $\mathcal{F} \| \mathcal{G}$.

Supongamos, sin pérdida de generalidad, que $\dim \mathcal{F} \leq \dim \mathcal{G}$. Probar:

$$\mathcal{F} \| \mathcal{G} \Longrightarrow (\mathcal{F} \cap \mathcal{G} = \varnothing \quad \text{ó} \quad \mathcal{F} \subseteq \mathcal{G}),$$

es equivalente a probar que:

$$\mathcal{F} \| \mathcal{G} \ \text{ y } \ \mathcal{F} \cap \mathcal{G} \neq \varnothing \Longrightarrow \mathcal{F} \subseteq \mathcal{G}.$$

Como $\mathcal{F} \| \mathcal{G}$, se tiene que $\boldsymbol{F} \subseteq \boldsymbol{G}$; y que $\mathcal{F} \cap \mathcal{G} \neq \varnothing$ (por la Proposición 2.21) implica que $\overrightarrow{PQ} \in \boldsymbol{F} + \boldsymbol{G}$. Se sigue, entonces que $\overrightarrow{PQ} \in \boldsymbol{G}$.

Luego, para probar que $\mathcal{F} \subseteq \mathcal{G}$, tomamos un punto arbitrario $X \in P + \boldsymbol{F}$:

$$X \in P + \boldsymbol{F} \Rightarrow \overrightarrow{PX} = \overrightarrow{PQ} + \overrightarrow{QX} \in \boldsymbol{F} \subseteq \boldsymbol{G} \Rightarrow \overrightarrow{QX} = \overrightarrow{PX} - \overrightarrow{PQ} \in \boldsymbol{G} \Rightarrow X \in Q + \boldsymbol{F}.$$

2.39 Ejemplo.- *Sea $\mathcal{F} = P + \boldsymbol{F}$ una variedad lineal de un espacio afín $\mathcal{A}, \dim \mathcal{F} = r$ y $Q \in \mathcal{A}$, entonces existe una única variedad lineal $\mathcal{F}$' de dimensión r, que pasa por Q y es paralela a $\mathcal{F}$.*

Basta tomar la variedad lineal $\mathcal{F}$' $= Q + \boldsymbol{F}$. Esta variedad lineal es única, ya que si existiera otra $\mathcal{F}$" $= Q + \boldsymbol{G}$ paralela a $\mathcal{F}$ y de dimensión r, se verificará que $\boldsymbol{F} = \boldsymbol{G}$. (Si $Q \in \mathcal{F}$, entonces $\mathcal{F}$' $= \mathcal{F}$).

2.40 Ejemplo.- *Dados una recta y un punto en un espacio afín, por el punto pasa una única recta paralela a la dada.*

2.41 Ejemplo.- *Dados un plano y un punto en un espacio afín, por el punto pasa uno y sólo uno paralelo al dado.*

TEMA 3
Coordenadas en un espacio afín. Ecuaciones de una variedad lineal

Un ente de gran utilidad en espacio vectoriales es una base, formada por vectores independientes; el correspondiente concepto en espacios afines es el de una referencia formada por un subconjunto de puntos, para lo cual necesitamos introducir la noción de dependencia e independencia de puntos.

Introduciremos dos tipos de referencias en un espacio afín: cartesiana y baricéntrica. La segunda es intrínseca en los espacio afín, en el sentido de que sólo involucra puntos; sin embargo, la primera necesita vectores del espacio vectorial al cual está asociado. Ambas referencias, nos permitirán introducir coordenadas y por tanto, a efectos prácticos, poder enfocar el estudio en la geometría afín manipulando relaciones algebraicas con escalares concretos. Por ejemplo, expresando la condición que un punto esté en una variedad lineal mediante el hecho de que sus coordenadas satisfagan a un sistema de ecuaciones lineales.

3.1 Coordenadas cartesianas en un espacio afín

Coordenadas cartesianas

3.1 Definición.- *Una* ***referencia cartesiana*** *en un espacio afín $(\mathcal{A}, \boldsymbol{E}, +)$ de $\dim \mathcal{A} = n$, consiste en fijar un punto $O \in \mathcal{A}$ (llamado origen) y una base $\{\vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$ de $\boldsymbol{E}$ y la indicamos por: $\mathcal{R} = \{O; \vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$.*
A las rectas que pasan por O y tienen la dirección del vector $\vec{\boldsymbol{e}}_k$, le denominamos k–ésimo eje de la referencia $\mathcal{R}$ o k–ésimo ***eje coordenado****.*
Llamamos ***coordenadas cartesianas*** *de un punto $X \in \mathcal{A}$ en la referencia $\mathcal{R}$ a las componentes $(x^1, \dots, x^n)$ de $\overrightarrow{OX}$ en la base $\{\vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$.*

Por notación, para indicar que $(x^1, \dots, x^n)$ son las coordenadas cartesianas de un punto X, escribiremos $X(x^1, \dots, x^n)$.

Conocidas las coordenadas cartesianas, en la referencia $\{O; \vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$, de dos puntos $P(x^1, \dots, x^n)$ y $Q(y^1, \dots, y^n)$, las componentes del vector $\overrightarrow{PQ}$ respecto de la base $\{\vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$ son $(y^1 - x^1, \dots, y^n - x^n)$, ya que $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = (y^1 - x^1)\vec{\boldsymbol{e}}_1 + \cdots + (y^n - x^n)\vec{\boldsymbol{e}}_n$.

Vamos a obtener las expresiones del cambio de coordenadas respecto a dos referencias cartesianas.

Sean $\mathcal{R} = \{P; \vec{\boldsymbol{u}}_1, \ldots, \vec{\boldsymbol{u}}_n\}$ y $\bar{\mathcal{R}} = \{Q; \vec{\boldsymbol{v}}_1, \ldots, \vec{\boldsymbol{v}}_n\}$ dos referencias cartesianas en un espacio afín $\mathcal{A}$ y sean $(q^1, \ldots, q^n)$ las coordenadas de Q respecto a $\mathcal{R}$ y $M = (a_i^j)$ la matriz cuyas columnas son las componentes de los vectores $\vec{\boldsymbol{v}}_i$ respecto a la base $\{\vec{\boldsymbol{u}}_1, \ldots, \vec{\boldsymbol{u}}_n\}$, es decir:

$$\overrightarrow{PQ} = q^1\vec{\boldsymbol{u}}_1 + \cdots + q^n\vec{\boldsymbol{u}}_n, \qquad \vec{\boldsymbol{v}}_i = \sum_{j=1}^{n} a_i^j \vec{\boldsymbol{u}}_j \quad (i = 1, \ldots, n).$$

Tomemos un punto $X \in \mathcal{A}$ con coordenadas cartesianas $(x^1, \ldots, x^n)$ y $(\bar{x}^1, \ldots, \bar{x}^n)$ en las referencias $\mathcal{R}$ y $\bar{\mathcal{R}}$, respectivamente, esto es:

$$\overrightarrow{PX} = x^1\vec{\boldsymbol{u}}_1 + \cdots + x^n\vec{\boldsymbol{u}}_n, \qquad \overrightarrow{QX} = \bar{x}^1\vec{\boldsymbol{v}}_1 + \cdots + \bar{x}^n\vec{\boldsymbol{v}}_n.$$

Entonces,

$$\overrightarrow{QX} = \sum_{i=1}^{n} \bar{x}^i\vec{\boldsymbol{v}}_i = \sum_{i=1}^{n} \bar{x}^i\Big(\sum_{j=1}^{n} a_i^j\vec{\boldsymbol{u}}_j\Big) = \sum_{j=1}^{n}\Big(\sum_{i=1}^{n} a_i^j\bar{x}^i\Big)\vec{\boldsymbol{u}}_j.$$

Como $\overrightarrow{PX} = \overrightarrow{PQ} + \overrightarrow{QX}$, se tiene:

$$x^j = q^j + \sum_{i=1}^{n} a_i^j \bar{x}^i \qquad (j = 1, \ldots, n).$$

En notación matricial la relación entre las coordenadas cartesianas del punto X, respecto a las dos referencias, se expresa por:

$$\begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} q^1 \\ q^2 \\ \vdots \\ q^n \end{pmatrix} + \begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \cdots & a_n^n \end{pmatrix} \begin{pmatrix} \bar{x}^1 \\ \bar{x}^2 \\ \vdots \\ \bar{x}^n \end{pmatrix}.$$

Las ecuaciones obtenidas nos dan las coordenadas cartesianas $(x^1, \ldots, x^n)$ de un punto X respecto a la referencia $\mathcal{R}$, conocidas sus coordenadas cartesianas $(\bar{x}^1, \ldots, \bar{x}^n)$ respecto a la referencia $\bar{\mathcal{R}}$, las componentes $(a_i^1, \ldots, a_i^n)$ de cada vector $\vec{\boldsymbol{v}}_i$ $(i = 1, \ldots, n)$, de la base $\{\vec{\boldsymbol{v}}_1, \ldots, \vec{\boldsymbol{v}}_n\}$, respecto a los de la base $\{\vec{\boldsymbol{u}}_1, \ldots, \vec{\boldsymbol{u}}_n\}$ y las coordenadas cartesianas $(q^1, \ldots, q^n)$ del punto Q respecto a la referencia $\mathcal{R}$.

En notación simbólica, podemos poner las ecuaciones cambio de referencias por

$$\begin{pmatrix} 1 \\ X \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ Q & M \end{pmatrix} \begin{pmatrix} 1 \\ \bar{X} \end{pmatrix},$$

siendo X y $\overline{X}$ matrices columna formadas por las coordenadas cartesianas $(x^1, \ldots, x^n)$ y $(\bar{x}^1, \ldots, \bar{x}^n)$ de un punto X respecto a las referencias $\mathcal{R}$ y $\bar{\mathcal{R}}$, respectivamente; Q la matriz columna formada por las coordenadas cartesianas $(q^1, \ldots, q^n)$ del punto Q respecto a la referencia $\mathcal{R}$; $M = (a_i^j)$ la matriz cambio de bases de $\{\vec{\boldsymbol{u}}_1, \ldots, \vec{\boldsymbol{u}}_n\}$ a $\{\vec{\boldsymbol{v}}_1, \ldots, \vec{\boldsymbol{v}}_n\}$; y donde 0 indica un matriz fila con n

ceros:

$$\begin{pmatrix} 1 \\ x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ q^1 & a_1^1 & a_2^1 & \cdots & a_n^1 \\ q^2 & a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q^n & a_1^n & a_2^n & \cdots & a_n^n \end{pmatrix} \begin{pmatrix} 1 \\ \bar{x}^1 \\ \bar{x}^2 \\ \vdots \\ \bar{x}^n \end{pmatrix}.$$

Ecuaciones paramétricas de una variedad lineal en coordenadas cartesianas

Sean $(\mathcal{A}, \boldsymbol{E}, +)$ un espacio afín de dimensión n, $\mathcal{R} = \{O; \vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$ una referencia cartesiana y $\mathcal{F}$ una variedad lineal contenida en $\mathcal{A}$, de subespacio director $\boldsymbol{F} \dot{\subset} \boldsymbol{E}$ y que pasa por un punto $P \in \mathcal{A}$ ($\mathcal{F} = P + \boldsymbol{F}$).

Supongamos que $\dim \boldsymbol{F} = r$, que $\{\vec{\boldsymbol{u}}_1, \dots \vec{\boldsymbol{u}}_r\}$ es una base de $\boldsymbol{F}$, vectores independientes, y

$$\vec{\boldsymbol{u}}_j = \sum_{i=1}^{n} a_j^i \vec{\boldsymbol{e}}_i \qquad (j = 1, \dots, r).$$

Si $X \in \mathcal{F}$, existe un único $\vec{\boldsymbol{v}} \in \boldsymbol{F}$, tal que $X = P + \vec{\boldsymbol{v}}$. Sean $(x^1, \dots, x^n)$ y $(p^1, \dots, p^n)$ las coordenadas cartesianas de X y P, respectivamente, en la referencia $\mathcal{R}$, y $(\lambda^1, \dots, \lambda^r)$ las componentes de $\vec{\boldsymbol{v}}$ en la base tomada en $\boldsymbol{F}$, es decir, $\vec{\boldsymbol{v}} = \sum_{j=1}^{r} \lambda^j \vec{\boldsymbol{u}}_j$, entonces la variedad lineal $\mathcal{F}$ es el conjunto de puntos cuyas coordenadas cartesianas en la referencia $\mathcal{R}$ se ponen en la forma:

$$\begin{array}{c} x^1 = p^1 + \lambda^1 a_1^1 + \lambda^2 a_2^1 + \cdots + \lambda^r a_r^1 \\ x^2 = p^2 + \lambda^1 a_1^2 + \lambda^2 a_2^2 + \cdots + \lambda^r a_r^2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x^n = p^n + \lambda^1 a_1^n + \lambda^2 a_2^n + \cdots + \lambda^r a_r^n \end{array} \tag{3-1}$$

Estas relaciones son las ***ecuaciones paramétricas*** de la variedad lineal $\mathcal{F}$, respecto a la referencia $\mathcal{R}$.

3.2 Nota.- Los escalares $(\lambda^1, \dots, \lambda^r)$, de las ecuaciones (3-1), son las coordenadas cartesianas de los puntos de la variedad lineal $\mathcal{F}$ (espacio afín $(\mathcal{F}, \boldsymbol{F}, +)$), respecto a la referencia cartesiana $\{P; \vec{\boldsymbol{u}}_1, \dots, \vec{\boldsymbol{u}}_r\}$ en $\mathcal{F}$.

En el caso de que se tenga un sistema de generadores $\{\vec{\boldsymbol{v}}_1, \dots, \vec{\boldsymbol{v}}_m\}$ ($m \geq r = \dim \mathcal{F}$) del subespacio vectorial $\boldsymbol{F}$ con

$$\vec{\boldsymbol{v}}_j = \sum_{i=1}^{n} b_j^i \vec{\boldsymbol{e}}_i \qquad (j = 1, \dots, m),$$

un punto $X = P + \vec{\boldsymbol{v}} \in \mathcal{F}$ se expresa en coordenadas por:

$$\begin{array}{c} x^1 = p^1 + \mu^1 b_1^1 + \mu^2 b_2^1 + \cdots + \mu^m b_m^1 \\ x^2 = p^2 + \mu^1 b_1^2 + \mu^2 b_2^2 + \cdots + \mu^m b_m^2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x^n = p^n + \mu^1 b_1^n + \mu^2 b_2^n + \cdots + \mu^m b_m^n \end{array} \tag{3-2}$$

siendo $\vec{\boldsymbol{v}} = \sum_{j=1}^{m} \mu^j \vec{\boldsymbol{v}}_j$. En esta circunstancias, no hay una correspondencia biyectiva entre las m–uplas $(\mu^1, \ldots, \mu^m)$ y los puntos de $\mathcal{F}$. Pero sí debe verificarse que el rango$(b^1_j) = r = \dim \boldsymbol{F}$.

3.3 Ejemplo.- *Las ecuaciones paramétricas de una recta que pasa por un punto P, de coordenadas cartesianas $(p^1, \ldots, p^n)$, y de subespacio director generado por $\vec{\boldsymbol{v}}$ (vector director de la recta), de componentes $(v^1, \ldots, v^n)$, son:*

$$x^1 = p^1 + \lambda v^1, \quad x^2 = p^2 + \lambda v^2, \ldots, x^n = p^n + \lambda v^n.$$

Que podemos escribir, eliminando el parámetro, en la forma:

$$\frac{x^1 - p^1}{v^1} = \frac{x^2 - p^2}{v^2} = \cdots = \frac{x^n - p^n}{v^n},$$

*y se conoce como **ecuación continua** de la recta.*

Ecuaciones cartesianas de una variedad lineal

Las ecuaciones paramétricas de un variedad lineal de dimensión r (3-1), en las incógnitas $\lambda^1, \ldots, \lambda^r$, han de tener solución única, para lo cual el rango de la matriz de coeficientes, rango$(a^i_j) = r$, ha de coincidir con el de la matriz ampliada con los términos independientes (de n filas y $r+1$ columnas); por lo que, una vez tomado un menor de orden r distinto de cero, en esta matriz, cualquier otro de orden $r+1$ ha de ser nulo. Esto da lugar a un sistema de $n-r$ ecuaciones lineales en las variables $x^1, \ldots, x^n$:

$$\begin{array}{l} \alpha^1_1 x^1 + \ldots\ldots\ldots + \alpha^1_n x^n = \beta^1 \\ \alpha^2_1 x^1 + \ldots\ldots\ldots + \alpha^2_n x^n = \beta^2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha^{n-r}_1 x^1 + \ldots + \alpha^{n-r}_n x^n = \beta^{n-r} \end{array} \tag{3-3}$$

Estas son las ***ecuaciones implícitas*** o ***ecuaciones cartesianas*** de la variedad lineal $\mathcal{F}$.

3.4 Ejemplo.- *La ecuación cartesiana de un hiperplano (variedad lineal de dimensión $n-1$ en un espacio afín de dimension n) es*

$$\alpha_1 x^1 + \cdots + \alpha_n x^n = \beta.$$

En particular, la ecuación cartesiana de una recta en un espacio afín bidimensional es:

$$ax + by + c = 0,$$

donde (x, y) son las coordenadas cartesianas de puntos del plano afín.

En el espacio afín tridimensional, la ecuación de un plano es:

$$ax + by + cz + d = 0,$$

donde (x, y, z) son las coordenadas cartesianas de puntos del espacio afín.

3.5 Nota.- Como generalización del Ejemplo 2.8 a un espacio afín cualquiera, un sistema de ecuaciones lineales $MX = B$ son las ecuaciones de una variedad lineal, donde X representa una matriz columna formada por las variables $(x^1, \dots, x^n)$, coordenadas cartesianas de un punto X de un espacio afín $\mathcal{A}$ de dimensión n, M es una matriz de m filas y n y B una matriz columna, ambas con coeficientes en K. Su dimension es $n - \operatorname{rango} M$.

Coordenadas homogéneas respecto a una referencia cartesiana

Consideremos en un espacio afín $(\mathcal{A}, \boldsymbol{E}, +)$ de dimensión n, una referencia cartesiana $\mathcal{R} = \{O; \vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$.

3.6 Definición.- *Sea $X \in \mathcal{A}$, con $X(x^1, \dots, x^n)$ sus coordenadas cartesianas respecto a la referencia $\mathcal{R}$, se definen las* `coordenadas homogéneas` *de X, respecto a $\mathcal{R}$, a la $n+1$–upla de escalares $(\xi^0, \xi^1, \dots, \xi^n)$ tales que*

$$\xi^0 \neq 0 \qquad x^1 = \frac{\xi^1}{\xi^0} \quad \dots \quad x^n = \frac{\xi^n}{\xi^0}. \tag{3-4}$$

Con esta condiciones, es claro que si $(\xi^0, \xi^1, \dots, \xi^n)$ son las coordenadas homogéneas de un punto X, también lo son $(\lambda\xi^0, \lambda\xi^1, \dots, \lambda\xi^n)$, para $\lambda \neq 0_K$; por lo que, denotamos las coordenadas homogéneas de un punto X, respecto a una referencia $\mathcal{R}$, por:

$$X(\xi^0 : \xi^1 : \cdots : \xi^n),$$

para indicar que se pueden tomar como coordenadas homogéneas del punto X esta $(n+1)$–upla o cualquier otra $(n+1)$–upla proporcional.

Existe un representante canónico de las coordenadas homogéneas de un punto X de coordenadas cartesianas $(x^1, \dots, x^n)$, que es

$$(1_K : x^1 : \cdots : x^n)$$

Si $(x^0 : x^1 : \cdots : x^n)$ son las coordenadas homogéneas de un punto, sus coordenadas cartesianas son $\left(\dfrac{x^1}{x^0}, \dots, \dfrac{x^n}{x^0}\right)$.

En el plano afín, si un punto tiene coordenadas cartesianas (x, y), denotamos sus coordenadas homogéneas por $(x^0 : x^1 : x^2)$, con $x^0 \neq 0, x = x^1/x^0$ e $y = x^2/x^0$.

En el espacio afín tridimensional, si un punto tiene coordenadas cartesianas (x, y, z), denotamos por $(x^0 : x^1 : x^2 : x^3)$ sus coordenadas homogéneas, con $x^0 \neq 0, x = x^1/x^0, y = x^2/x^0$ y $z = x^3/x^0$.

Ecuaciones homogéneas de una variedad lineal

A partir de las ecuaciones cartesianas (3-3) de una variedad podemos obtener sus ecuaciones homogéneas, usando la relación (3-4) que existe entre las coordenadas cartesianas y homogéneas:

$$\begin{array}{ccccccccc} \alpha_0^1 x^0 & + & \alpha_1^1 x^1 & + & \cdots & + & \alpha_n^1 x^n & = & 0 \\ \alpha_0^2 x^0 & + & \alpha_1^2 x^1 & + & \cdots & + & \alpha_n^2 x^n & = & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ \alpha_0^{n-r} x^0 & + & \alpha_1^{n-r} x^1 & + & \cdots & + & \alpha_n^{n-r} x^n & = & 0, \end{array} \tag{3-5}$$

donde $(x^0 : x^1 : \cdots : x^n)$ son las coordenadas homogéneas deducidas de una referencia cartesiana.

Puntos del infinito asociados a una referencia cartesiana en el espacio afín real

Sea $(\mathcal{A}, \mathbb{R}^n, +)$ un espacio afín real de dimensión n y $\mathcal{R} = \{O; \vec{e}_1, \dots, \vec{e}_n\}$ una referencia cartesiana: todo punto $X \in \mathcal{A}$ tiene unas coordenadas cartesianas $(x^1, \dots, x^n)$ respecto a esta referencia, y unas coordenadas homogéneas, deducidas de ellas, $(\xi^0, \xi^1, \dots, \xi^n)$, $\xi^0 \neq 0$. Consideremos la recta $\mathcal{L}$ que pasa por un punto $P(p^1, \dots, p^n)$ y de dirección la determinada por el vector $\vec{v} \neq \vec{o}$, de componentes $(v^1, \dots, v^n)$ respecto a la base $\{\vec{e}_1, \dots, \vec{e}_n\}$. Las ecuaciones paramétricas de $\mathcal{L}$ son:

$$x^1 = p^1 + \lambda v^1, x^2 = p^2 + \lambda v^2, \dots, x^n = p^n + \lambda v^n.$$

Si recorremos la recta $\mathcal{L}$ a partir del punto P, alejándonos indefinidamente, para valores de λ tendiendo a infinito, obtendremos una n–upla con alguna de sus componentes que no son números reales. Se trata de solucionar el problema de asignar al punto del infinito unas coordenadas reales; lo que podemos solucionar acudiendo a las coordenadas homogéneas.

Un punto genérico de la recta $\mathcal{L}$, distinto de P, tiene por coordenadas homogéneas, para $\lambda \neq 0$:

$$\left(1 : p^1 + \lambda v^1 : p^2 + \lambda v^2 : \cdots : p^n + \lambda v^n\right) = \left(\frac{1}{\lambda} : \frac{p^1}{\lambda} + v^1 : \frac{p^2}{\lambda} + v^2 : \cdots : \frac{p^n}{\lambda} + v^n\right).$$

Que cuando $\lambda \to \infty$, tiende a la $(n+1)$–upla $(0 : v^1 : v^2 : \cdots : v^n)$ de números reales, que no constituyen las coordenadas homogéneas de ningún punto de espacio afín $\mathcal{A}$, pues la primera componente es nula.

Añadiremos un punto a los de $\mathcal{A}$, llamado ***punto del infinito*** de la recta $\mathcal{L}$ cuyas coordenadas homogéneas, respecto a la referencia $\mathcal{R}$, son:

$$(0 : v^1 : v^2 : \cdots : v^n),$$

3.7 Nota.- Las coordenadas homogéneas $(0 : v^1 : v^2 : \cdots : v^n)$ del punto del infinito de la recta $\mathcal{L}$ coinciden con las del punto del infinito de cualquier recta paralela a ella. Ya que otra recta paralela a $\mathcal{L}$ tiene de vector director el vector $\lambda\vec{v}$ $(\lambda \neq 0)$.

La ecuación de un hiperplano (variedad lineal de dimensión $n-1$) tiene por ecuación en coordenadas homogéneas:

$$a_0x^0 + a_1x^1 + \cdots + a_nx^n = 0,$$

con $a_0, a_1, \ldots, a_n$ números reales. Como los puntos del infinito de las rectas del espacio afín real, verifican la ecuación:

$$x^0 = 0,$$

se dice que constituyen el ***hiperplano del infinito***.

Pertenecen al hiperplano del infinito los puntos del infinito de las rectas que pasan por el origen de coordenadas O y tienen la dirección de los vectores que forman parte de la referencia cartesiana $\{O; \vec{e}_1, \ldots, \vec{e}_n\}$. Ellos tiene por coordenadas homogéneas:

$$(0:1:0:\cdots:0),\ (0:0:1:\cdots,0),\ \ldots\ ,(0:\cdots:0:1).$$

En el caso del plano afín real tridimensional con coordenadas de sus puntos (x, y), respecto a una referencia cartesiana, y con coordenadas homogéneas asociadas (x^0, x^1, x^2), la ***recta del infinito***, tiene por ecuación:

$$x^0 = 0.$$

En el espacio afín real con coordenadas de sus puntos (x, y, z), respecto a una referencia cartesiana, y con coordenadas homogéneas asociadas (x^0, x^1, x^2, x^3), el ***plano del infinito***, tiene por ecuación:

$$x^0 = 0.$$

3.2 Coordenadas baricéntricas en un espacio afín

Las coordenadas cartesianas que se han introducido en §3.1 están definidas partiendo de antemano de una base del espacio vectorial al que el espacio afín está asociado; lo más natural es que se parta de ciertos puntos en el espacio afín para definir coordenadas, lo cual tiene más carácter intrínseco. Necesitamos primeramente introducir los conceptos de dependencia e independencia de puntos.

Dependencia (afín) de puntos

3.8 Definición.- *Los puntos $P_1, P_2, \ldots, P_m$ de un espacio afín $\mathcal{A}$, asociado a un espacio vectorial $\mathbf{E}$, se dice que son* ***linealmente independientes*** *si los vectores $\overrightarrow{P_iP_1}, \ldots, \overrightarrow{P_iP_{i-1}}, \overrightarrow{P_iP_{i+1}}, \ldots, \overrightarrow{P_iP_m}$ son linealmente independientes en $\mathbf{E}$, para algún $i \in \{1, 2, \ldots, m\}$.*

Para que esta definición sea consistente, no puede depender del punto P_i elegido. Esto queda resuelto en la proposición siguiente:

3.9 Proposición.- *Dado un espacio afín $(\mathcal{A}, \mathbf{E}, +)$, sea $\{P_1, \ldots, P_m\} \subset \mathcal{A}$. Si $\exists i \in \{1, 2, \ldots, m\}$ tal que $\{\overrightarrow{P_iP_j}\}_{j=1,2,\ldots,i-1,i+1,\ldots,m}$ es un conjunto de vectores independientes, entonces*

$$\forall k \in \{1, 2, \ldots, m\} : \{\overrightarrow{P_kP_j}\}_{j=1,2,\ldots,k-1,k+1,\ldots,m}$$

es un conjunto de vectores independientes.

Demostración.- Sin pérdida de generalidad, podemos tomar $i = 1$ (reordenando los puntos si es necesario) y que $\{\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}, \ldots, \overrightarrow{P_1P_m}\}$ sea un conjunto de vectores linealmente independientes.

Sea $k \in \{2, 3, \ldots, m\}$ y supongamos que existan $\{\lambda^j\}_{j=1,2,\ldots,k-1,k+1,\ldots,m}$ escalares tales que:

$$\sum_{\substack{j=1 \\ (j\neq k)}}^{m} \lambda^j \overrightarrow{P_kP_j} = \vec{\boldsymbol{0}}.$$

Ya que $\overrightarrow{P_kP_j} = \overrightarrow{P_kP_1} + \overrightarrow{P_1P_j}$,

$$\begin{aligned} \sum_{\substack{j=1 \\ (j\neq k)}}^{m} \lambda^j \overrightarrow{P_kP_j} &= \sum_{\substack{j=1 \\ (j\neq k)}}^{m} \lambda^j \overrightarrow{P_kP_1} + \sum_{\substack{j=1 \\ (j\neq k)}}^{m} \lambda^j \overrightarrow{P_1P_j} = \\ &= -\Big(\sum_{\substack{j=1 \\ (j\neq k)}}^{m} \lambda^j\Big)\overrightarrow{P_1P_k} + \sum_{\substack{j=2 \\ (j\neq k)}}^{m} \lambda^j \overrightarrow{P_1P_j} = \vec{\boldsymbol{0}}. \end{aligned}$$

Como $\{\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}, \ldots, \overrightarrow{P_1P_m}\}$ es un conjunto de vectores linealmente independientes:

$$\lambda^2 = \cdots = \lambda^{k-1} = -\big(\sum_{\substack{j=1 \\ (j\neq k)}}^{m} \lambda^j\big) = \lambda^{k+1} = \cdots = \lambda^m = 0_K.$$

De donde se deduce que también $\lambda^1 = 0_K$. ◾

El siguiente resultado nos da otra caracterización de puntos independientes.

3.10 Proposición.- *Dado un espacio afín $(\mathcal{A}, \mathbf{E}, +)$, $\{P_1, \ldots, P_m\} \subset \mathcal{A}$ es un conjunto de puntos linealmente independientes si y sólo si*

$$\forall P \in \mathcal{A}, \sum_{i=1}^{m} \lambda^i \overrightarrow{PP_i} = \vec{\boldsymbol{0}}, \lambda^1 + \cdots + \lambda^m = 0_K \Rightarrow \lambda^1 = \cdots = \lambda^m = 0_K.$$

Demostración.- Supongamos que $P_1, \ldots, P_m$ son linealmente independientes, o sea que $\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}, \ldots, \overrightarrow{P_1P_m}$ son linealmente independientes.

Si $\lambda^1 \overrightarrow{PP_1} + \cdots + \lambda^m \overrightarrow{PP_m} = \vec{\boldsymbol{0}}$ y $\lambda^1 + \cdots + \lambda^m = 0_K$, se tiene que:

$$\begin{aligned} &(-\lambda^2 - \cdots - \lambda^m)\overrightarrow{PP_1} + \lambda^2 \overrightarrow{PP_2} + \cdots + \lambda^m \overrightarrow{PP_m} = \\ &= \lambda^2(\overrightarrow{PP_2} - \overrightarrow{PP_1}) + \cdots + \lambda^m(\overrightarrow{PP_m} - \overrightarrow{PP_1}) = \end{aligned}$$

$$\left.\begin{array}{c} = \lambda^2 \overrightarrow{P_1P_2} + \cdots + \lambda^m \overrightarrow{P_1P_m} = \vec{\boldsymbol{0}} \\ \overrightarrow{P_1P_2}, \ldots, \overrightarrow{P_1P_m} \text{ independientes} \\ \lambda^1 + \cdots + \lambda^m = 0_K \end{array}\right\} \Rightarrow \lambda^1 = \lambda^2 = \cdots = \lambda^m = 0_K.$$

Recíprocamente, partimos de $\lambda^2\overrightarrow{P_1P_2}+\cdots+\lambda^m\overrightarrow{P_1P_m} = \vec{\boldsymbol{0}}$ y vamos a deducir que todos los escales son nulos; con lo que los vectores $\overrightarrow{P_1P_2},\ldots,\overrightarrow{P_1P_m}$ son linealmente independientes; es decir, los puntos $P_1, P_2, \ldots, P_m$ son linealmente independientes.

Para un punto arbitrario $P \in \mathcal{A}$:

$\sum\limits_{i=2}^{m} \lambda^i\overrightarrow{P_1P_i} = \sum\limits_{i=2}^{m} \lambda^i(\overrightarrow{PP_i} - \overrightarrow{PP_1}) = -\big(\sum\limits_{i=2}^{m} \lambda^i\big)\overrightarrow{PP_1} + \sum\limits_{i=2}^{m} \lambda^i\overrightarrow{PP_i} = \vec{\boldsymbol{0}}.$

Como se verifica que la suma de los coeficientes que aparecen en esta combinación lineal es cero, todos son nulos; es decir, $\lambda^2 = \cdots = \lambda^m = 0_K$. ◪

3.11 Nota.- En un espacio afín de dimensión n, el número máximo de puntos linealmente independientes es $n+1$.

3.12 Nota.- Una propiedad notable de la independencia de vectores $\vec{\boldsymbol{u}}_1,\ldots,\vec{\boldsymbol{u}}_m$ de un espacio vectorial es que si $\vec{\boldsymbol{v}}$ es combinación lineal de ellos $\vec{\boldsymbol{v}} = \sum\limits_{k=1}^{m} \lambda^k u_k$, entonces los escalares $\lambda^1,\ldots,\lambda^m$ son únicos. Un resultado similar se tiene para puntos independientes en un espacio afín (con notaciones que introduciremos posteriormente):

En un espacio afín $(\mathcal{A}, \mathbf{E}, +)$ se dan los $m+1$ puntos $P_0, P_1, \ldots, P_m$, sea $X \in \mathcal{A}$ y supongamos que $X = \sum\limits_{k=0}^{m} \lambda^k P_k$, con $\sum\limits_{k=1}^{m} \lambda^k = 1$. Entonces, los escales $\lambda^0, \lambda^1, \ldots, \lambda^m$ son únicos si sólo si los puntos $P_0, P_1, \ldots, P_m$ son linealmente independientes

Baricentro

3.13 Definición.- *Se llama* **baricentro** *de los puntos $P_1,\ldots,P_m$ afectados de los "pesos" (las masas o los escalares o los coeficientes) $\alpha^1,\ldots,\alpha^m \in K$, tales que $\alpha^1+\cdots+\alpha^m \neq 0_K$, al punto B determinado por*

$$\overrightarrow{PB} = \frac{\alpha^1\overrightarrow{PP_1}+\cdots+\alpha^m\overrightarrow{PP_m}}{\alpha^1+\cdots+\alpha^m} = \frac{\sum\limits_{k=1}^{m} \alpha^k\overrightarrow{PP_k}}{\sum\limits_{k=1}^{m} \alpha^k},$$

donde P es un punto cualquiera de $\mathcal{A}$.

En esta definición no tiene relevancia el punto P tomado. En efecto, tomemos otro punto Q entonces si

$$\overrightarrow{QB'} = \frac{\sum\limits_{k=1}^{m} \alpha^k\overrightarrow{QP_k}}{\sum\limits_{k=1}^{m} \alpha^k},$$

se debe tener que $B = B$'.

Para $k = 1, \dots, m$ se tiene que $\overrightarrow{PQ} = \overrightarrow{PP_k} - \overrightarrow{QP_k}$ y además, $\overrightarrow{BB'} = \overrightarrow{BP} + \overrightarrow{PB'} = \overrightarrow{BP} + \overrightarrow{PQ} + \overrightarrow{QB'} = \overrightarrow{PQ} + \overrightarrow{QB'} - \overrightarrow{PB}$, entonces:

$$\overrightarrow{BB'} = \overrightarrow{PQ} + \frac{\sum_{k=1}^{m} \alpha^k(\overrightarrow{QP_k} - \overrightarrow{PP_k})}{\sum_{k=1}^{m} \alpha^k} = \overrightarrow{PQ} + \frac{(\sum_{k=1}^{m} \alpha^k)\overrightarrow{QP}}{\sum_{k=1}^{m} \alpha^k} = \overrightarrow{PQ} - \overrightarrow{PQ} = \vec{\boldsymbol{0}}.$$

Por tanto $\overrightarrow{BB}$'$= \vec{\boldsymbol{0}} \Rightarrow B = B$'.

Una primera propiedad del baricentro es la "homogeneidad": no cambia el baricentro si se multiplica todas los "pesos" por un mismo factor $\lambda \neq 0$.

Si los "pesos" $\alpha^1, \dots, \alpha^m$ que afectan a los puntos $P_1, \dots, P_m$, verifican que $\alpha^1 + \cdots + \alpha^m = 1_K$, su baricentro es el punto tal que

$$\overrightarrow{PB} = \alpha^1\overrightarrow{PP_1} + \cdots + \alpha^m\overrightarrow{PP_m},$$

y como no depende del punto P, se conviene en utilizar la siguiente notación para el baricentro B de $P_1, \dots, P_m$ afectados de las "pesos" $\alpha^1, \dots, \alpha^m$:

$$B = \alpha^1 P_1 + \cdots + \alpha^m P_m. \tag{3-6}$$

Tomando $B = P$ en la Definición 3.13, se obtiene esta nueva definición de baricentro:

3.13' Definición.- *Se llama* ***baricentro*** *de los puntos $P_1, \dots, P_m$ afectados de los "pesos" $\alpha^1, \dots, \alpha^m$, escalares tales que $\alpha^1 + \cdots + \alpha^m \neq 0_K$, al punto B determinado por*

$$\alpha^1\overrightarrow{BP_1} + \cdots + \alpha^m\overrightarrow{BP_m} = \vec{\boldsymbol{0}}.$$

3.14 Nota.- Puede ser conveniente, para tratar el concepto de baricentro, introducir la noción de puntos ponderados, como un par (P, α), $P \in \mathcal{A}$ y $\alpha \in K$; entonces, decir que B es el baricentro de la familia de puntos ponderados $\{(P_k, \alpha^k)\}_{k=1,\dots,m}$.

3.15 Nota.- En general, si no se especifican "pesos", se supone que todos son iguales. Y por tanto, B es el baricentro de $P_1, \dots, P_m$ si y sólo si

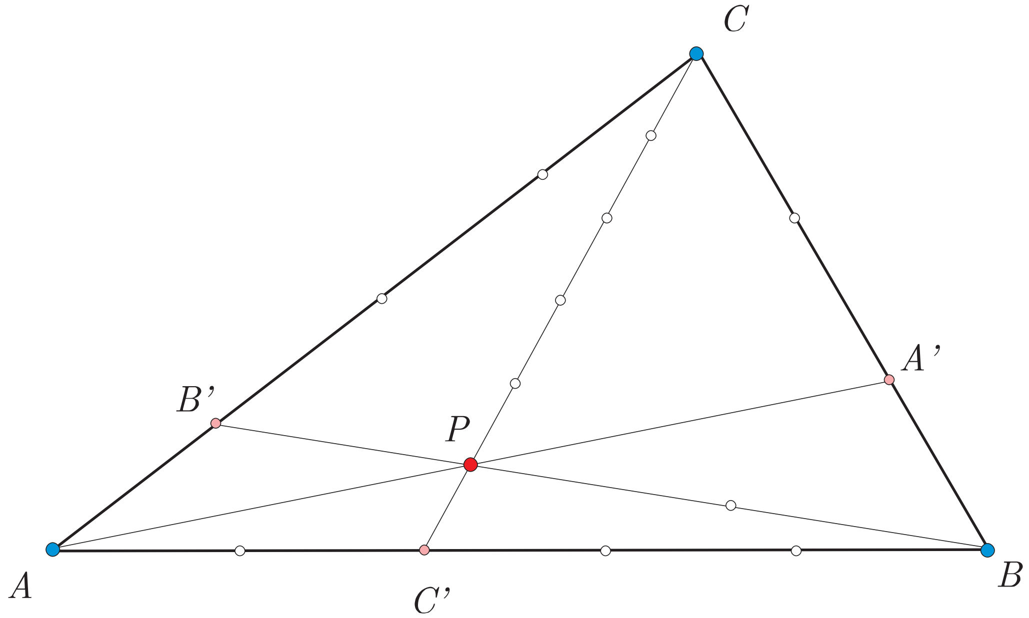
$$\sum_{k=1}^{m} \overrightarrow{BP_i} = \overrightarrow{BP_1} + \cdots + \overrightarrow{BP_m} = \vec{\boldsymbol{0}}.$$

Si $m = 2$, el baricentro se dice que es el ***punto medio*** de los punto P_1 y P_2. Para $m = 3$ se llama baricentro (o centroide) del triángulo $\widehat{P_1P_1P_3}$. En estos casos particulares el cuerpo K no puede ser de característica 2 o 3, respectivamente.

Otra propiedad del baricentro es la "asociatividad": el baricentro se puede calcular reagrupando puntos, es decir introduciendo baricentros parciales.

Por ejemplo, si D es baricentro de $\{(A,\alpha),(B,\beta)\}$ (con $\alpha+\beta\neq 0$) entonces, el baricentro de $\{(A,\alpha),(B,\beta),(C,\gamma)\}$ es el baricentro de $\{(D,\alpha+\beta),(C,\gamma)\}$ $(\alpha+\beta+\gamma\neq 0)$.

3.16 Ejemplo.- *En el plano afín real podemos ilustrar la construcción gráfica del baricentro P de los puntos ponderados $(A,3)$, $(B,2)$ y $(C,1)$, determinando primeramente el baricentro de los puntos ponderados $(B,2)$ y $(C,1)$; será un punto A' en la recta BC tal que $2\overrightarrow{A'B}+1\overrightarrow{A'C}=\vec{\boldsymbol{0}}$, o sea $\overrightarrow{A'C}=-2\overrightarrow{A'B}$, es decir, A' es el punto más cerca de B que resulta al trisecar el segmento BC. Luego, se determina el baricentro P de los puntos ponderados $(A,3)$ y $(A',2+1)$, que es el punto medio de A y A'.*



Applet-Cabri *http://webpages.ull.es/users/amontes/cabri/g1-07.htm*

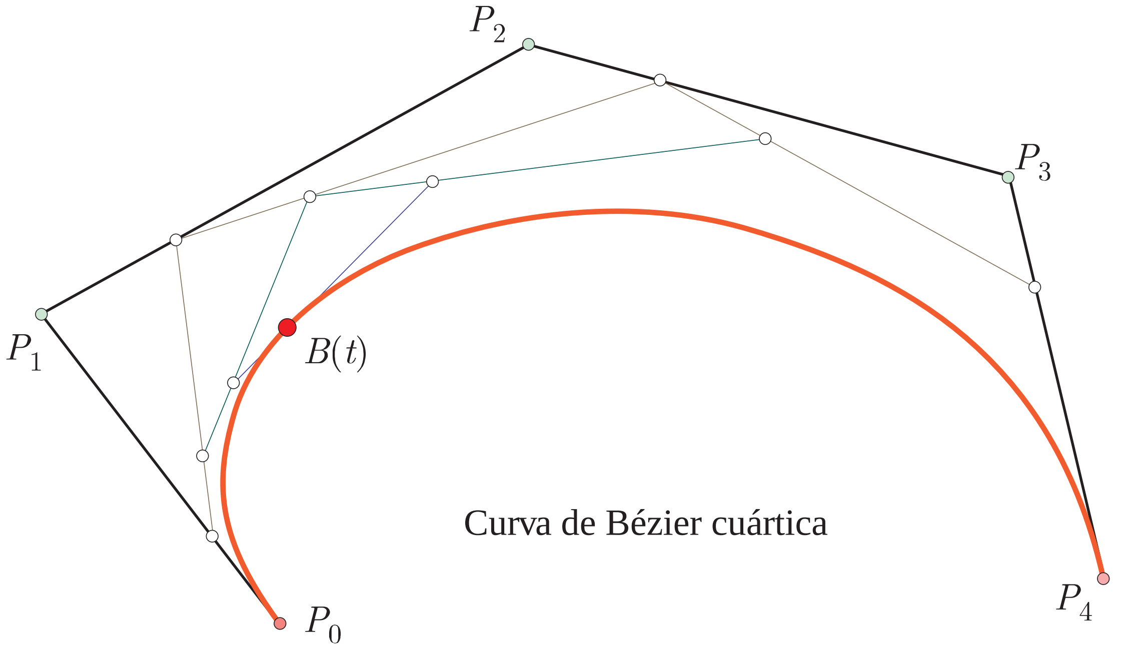
Para determinar el baricentro P de los puntos ponderados $(A,3),(B,2)$ y $(C,1)$, podemos empezar con cualquier par de puntos, y el resultado final será el mismos punto; así, por ejemplo, sea B' el baricentro de $(A,3)$ y $(C,1)$, que será el punto más cerca de A de la división del segmento AC en cuatro partes iguales; luego, hallar el baricentro P de $(B,2)$ y $(B',3+1)$, que es el punto más cerca de B' que resulta al trisecar el segmento BC.

Una situación más general que la expuesta, relativa a la asociatividad del baricentro, es el siguiente resultado:

3.17 Proposición.- *Sean $P_1,\ldots,P_m$, m puntos en un espacio afín $\mathcal{A}$ y sean $\alpha_1,\ldots,\alpha_m$ sus "pesos". Supongamos que $I_1,\ldots,I_r$ son subconjuntos disjuntos de $\{1,\ldots,m\}$ cuya unión es $\{1,\ldots,m\}$. Sea B_j el baricentro de los puntos $\{P_k/k\in I_j\}$ con los "pesos" correspondientes. Entonces, el baricentro de los puntos $B_1,\ldots,B_r$ con "pesos" $\beta_j=\sum_{i\in I_j}\alpha_i$ es exactamente el baricentro de los puntos $P_1,\ldots,P_m$ afectados de los "pesos" $\alpha_1,\ldots,\alpha_m$.* ◼

3.18 Ejemplo.- *Curvas de Bézier*

En el plano afín real el baricentro de los puntos P_0, P_1, P_2, P_3 y P_4, afectados de los "pesos" $(1-t)^4, 4(1-t)^3t, 6(1-t)^2t^2, 4(1-t)t^3$ y t^4, respectivamente, es el punto P que recorre una cuártica, tangente en P_0 a la recta P_0P_1 y en el punto P_4 a la recta P_3P_4. Los puntos P_0 y P_4 son los nodos y los puntos P_1, P_2 y P_3 son los puntos de control de la curva de Bézier.



Applet-GeoGebra http://webpages.ull.es/users/amontes/geogebra/g1_06.html

Como $1 = ((1-t)+t)^4 = (1-t)^4 + 4(1-t)^3t + 6(1-t)^2t^2 + 4t^3(1-t) + t^4$, los "pesos" de los puntos P_0, P_1, P_2, P_3, P_4 suman 1, y su baricentro es, adoptando la notación (3-6):

$$B(t) = (1-t)^4P_0 + 4(1-t)^3tP_1 + 6(1-t)^2t^2P_2 + 4t^3(1-t)P_3 + t^4P_4.$$

Se denomina curvas de Bézier a un sistema que se desarrolló hacia los años 1960, para el trazado de dibujos técnicos, en el diseño aeronáutico y de automóviles. Su denominación es en honor a Pierre Bézier, quien ideó un método de descripción matemática de las curvas que se comenzó a utilizar con éxito en los programas de diseño asistido por computadora u ordenador, más conocido por sus siglas inglesas CAD (computer-aided design).

Las curvas de Bézier fueron publicadas, por primera vez en 1962 por el ingeniero de origen francés Pierre Bézier, que las usó en el diseño de las diferentes partes de la carrocería de un automóvil, en sus años de trabajo en la Renault. Las curvas fueron desarrolladas por Paul de Casteljau usando el algoritmo que lleva su nombre, para su representación gráfica.

Posteriormente, los inventores del PostScript [(1)], lenguaje que permitió el desarrollo de sistemas de impresión de alta calidad desde el ordenador, introdujeron en ese código el método de Bézier para la generación del código de las curvas y los trazados.

Curvas lineales de Bézier:

(1) El operador x_1 y_1 x_2 y_2 x_3 y_3 `curveto` de PostScript produce una cúbica de Bézier entre el punto actual (x_0, y_0) y el punto (x_3, y_3) usando los puntos (x_1, y_1) y (x_2, y_2) como puntos control. El siguiente código produce la cúbica de Bézier con nodos $(30, 100)$ y $(60, 120)$ y puntos de control $(50, 180)$ y $(100, 140)$:

Dados los puntos P_0 y P_1, una curva lineal de Bézier es el segmento de recta entre los dos puntos. La curva viene dada por la expresión:

$$B(t) = (1-t)P_0 + tP_1, \qquad t \in [0,1].$$

Curvas cuadráticas de Bézier:

Una curva cuadrática de Bézier es el camino trazado por la función $B(t)$, dados los puntos: P_0, P_1 y P_2,

$$B(t) = (1-t)^2 P_0 + 2(1-t)tP_1 + t^2 P_2, \qquad t \in [0,1].$$

Curvas cúbicas de Bézier

Cuatro puntos del plano P_0, P_1, P_2 y P_3 definen una curva cúbica de Bézier. La curva comienza en el punto P_0 su tangente en este punto tiene la dirección de la recta P_0P_1 y $\overrightarrow{P_0P_1}$ es el vector velocidad. Termina en el punto P_3 con vector velocidad en este punto (si se recorriera en sentido contrario) el vector $\overrightarrow{P_3P_2}$.

La forma paramétrica de la curva es:

$$B(t) = (1-t)^3 P_0 + 3(1-t)^2 tP_1 + 3(1-t)t^2 P_2 + t^3 P_3, \quad t \in [0,1].$$

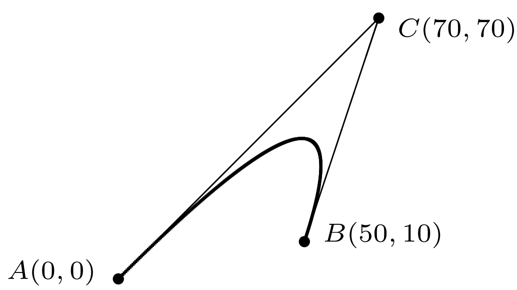
De forma general, una curva de Bézier de grado m se obtiene tomando como "pesos" de los $m+1$ punto $P_0, P_1, \ldots, P_m$ los coeficientes del polinomio $\binom{m}{k} t^{m-k}(1-t)^k$, $k = 0, 1, \ldots, m$:

$$B(t) = \binom{m}{0}(1-t)^m P_0 + \binom{m}{1}(1-t)t^{m-1} P_1 + \cdots + \binom{m}{m} t^m P_m.$$

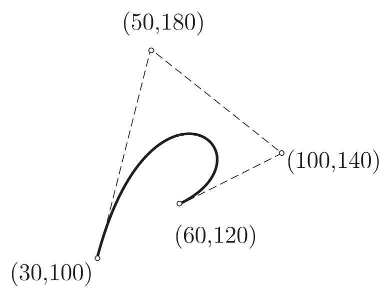
También son utilizadas las curvas Bézier en el sistema para la composición de textos científicos que incluyen muchos elementos de notación matemática como es el LaTeX, que está formado por un gran conjunto de macros de TeX, escrito por Leslie Lamport en 1984, con la intención de facilitar el uso del lenguaje de composición tipográfica, creado por Donald Knuth.

El comando `\qbezier` toma tres puntos como argumentos y dibuja una curva de Bézier cuadrática, tomado el primero y tercero como nodos y el segundo como punto control. Por ejemplo, el siguiente código genera una arco de parábola tangente en los $A(0,0)$ y $B(50,10)$ a las rectas AC y BC, respectivamente, siendo $C(70,70)$:

```
\begin{picture}(60,70)
\qbezier(0,0)(70,70)(50,10)
\end{picture}
```



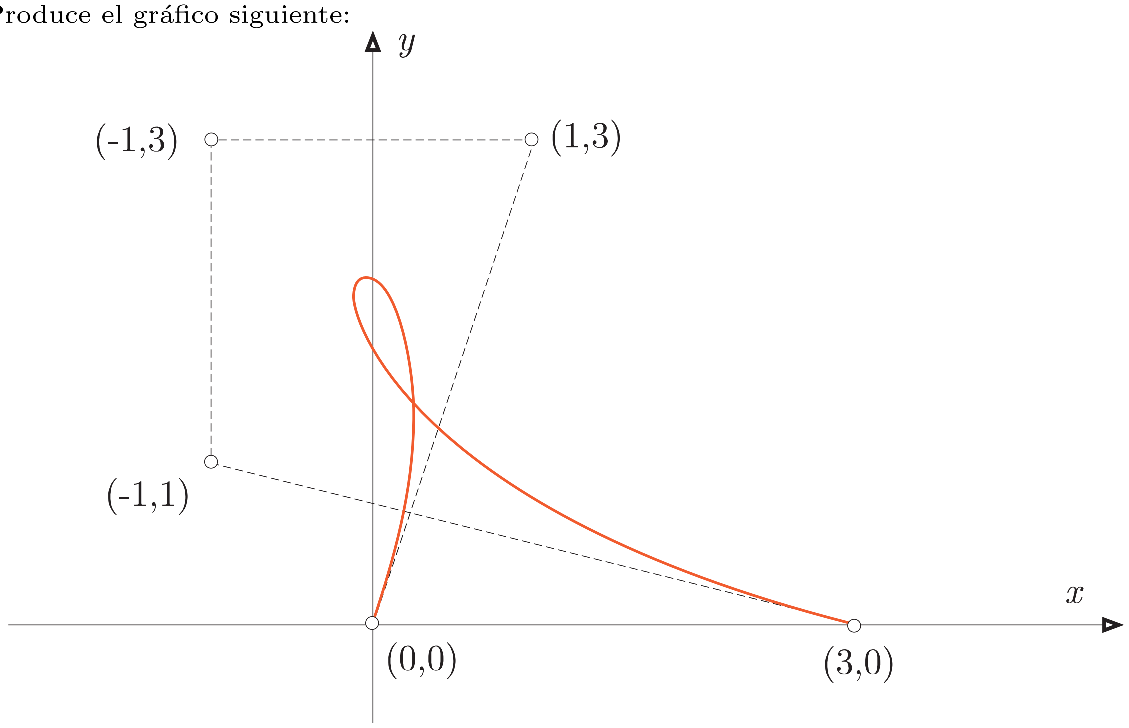

```
30 100 moveto
50 180 100 140 60 120 curveto
2 setlinewidth
stroke
showpage
```



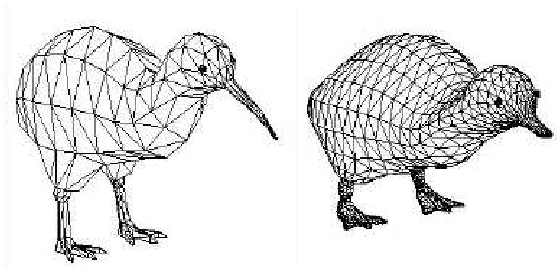
En MATHEMATICA de Wolfram Research, Inc. el comando **Spline** con la opción tipo **Bezier**, permite trazas curvas Bézier; por ejemplo:

```
<<Graphics`Spline`;
pts =  {{0, 0}, {1, 3}, {-1, 3}, {-1, 1}, {3, 0}};
Show[Graphics[{{Dashing[{0.05}], Line[pts]},
              Hue[0],Spline[pts,Bezier,SplineDots->True]}]]
```

Produce el gráfico siguiente:



Se obtendrán superficies de Bézier cuando las curvas paramétricas sean curvas de Bézier. Es el caso del siguiente gráfico tomado de Faramarz Samavati.- Modeling for Computer Graphics. University of Calgary.



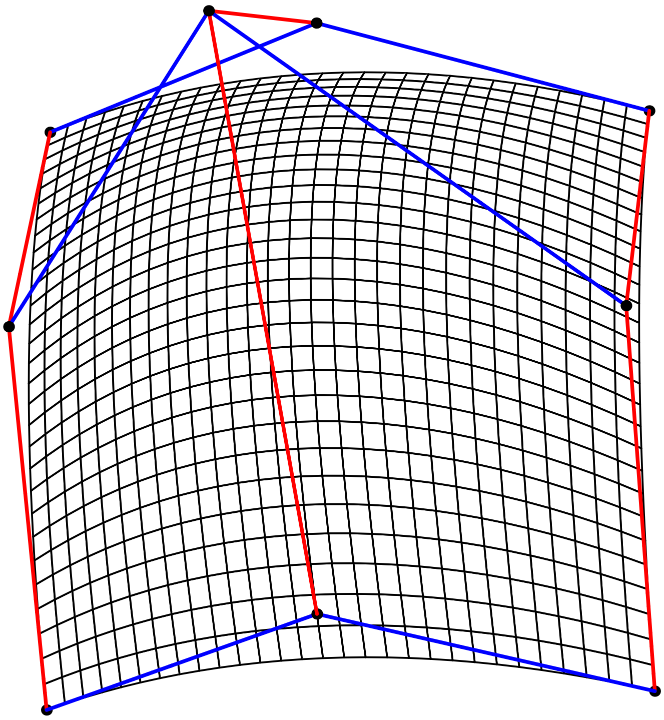
Otro ejemplo es el gráfico que se muestra a continuación que ha sido obtenido a partir de código em Mathematica siguiente, extraído de:

David Salomon. "Curves and Surfaces for Computer Graphics". Springer Verlag, 2005. (Figure 6.20, page 221, A Biquadratic Bézier Surface Patch).

```
(*  biquadratic bézier surface patch  *)

Clear[pwr, bern, spnts, hlines, vlines, n, m, bzSurf, g1, g2];
spnts = {{{0, 0, 0}, {1, 0, 1}, {0, 0, 2}},
         {{1, 1, 0}, {4, 1, 1}, {1, 1, 2}},
```

```
         {{0, 2, 0}, {1, 2, 1}, {0, 2, 2}}};

m = Length[spnts[[1]]] - 1; n = Length[Transpose[spnts][[1]]] - 1;

vlines = Graphics3D[{AbsoluteThickness[1],
        RGBColor[0, 0, 1],(*control polygon*)
        Table[Line[{spnts[[i, j]], spnts[[i + 1, j]]}], {i, 1, n}, {j, 1,m + 1}]}];

hlines = Graphics3D[{AbsoluteThickness[1], RGBColor[1, 0, 0],
        Table[Line[{spnts[[i, j]], spnts[[i, j + 1]]}], {i, 1, n + 1}, {j, 1,m}]}];

(*Handle Indeterminate condition*)
pwr[x_, y_] := If[x == 0 && y == 0, 1, x^y];

bern[n_, i_, u_] := Binomial[n, i]pwr[u, i]pwr[1 - u, n - i]
bzSurf[u_, w_] :=
  Sum[bern[n, i, u] spnts[[i + 1, j + 1]] bern[n, j, w], {i, 0, n}, {j, 0,
      n}]
g1 = ParametricPlot3D[bzSurf[u, w], {u, 0, 1}, {w, 0, 1}, Ticks -> False,
      Axes -> False, Compiled -> False, DisplayFunction -> Identity];
g2 = Graphics3D[{AbsolutePointSize[3],
        Table[Point[spnts[[i, j]]], {i, 1, n + 1}, {j, 1, n + 1}]}];

Show[g1, g2, vlines, hlines, ViewPoint -> {9, 1, -3}, PlotRange -> All,
    DefaultFont -> {"cmr10", 10}, Boxed -> False, Shading -> False,
    DisplayFunction -> $DisplayFunction];
```

JavaView http://webpages.ull.es/users/amontes/jview/g1-22.html



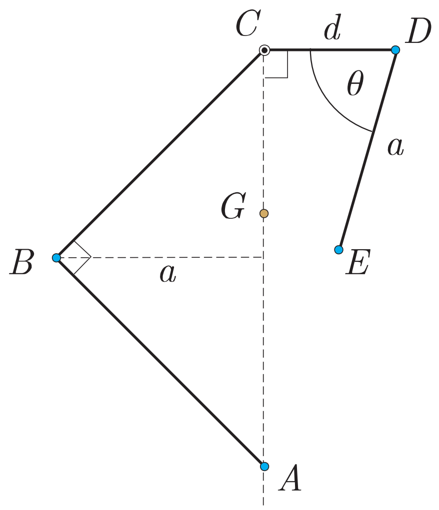
3.19 Ejemplo.- *Un alambre homogéneo ABCDE (situado todo él en un plano) está doblado como indica la figura y se sostiene con un pasador*

puesto en C, cumpliéndose que los puntos A, B y C forman un triángulo rectángulo isósceles, con el ángulo recto en B e hipotenusa $AC = 2a$; CA y CD son perpendiculares; la longitud del segmento CD es d; θ es el ángulo $\widehat{CDE}$; y la longitud del segmento $DE = a$.

Es necesario que $d \leq a$ para que el segmento CD del alambre puede ser horizontal.

La relación que existe entre a, d y θ para el segmento CD del alambre se mantenga horizontal es $2d = a(1 + \cos\theta)$.

Tomando un sistema de coordenadas cartesiano en el que $C(0,0)$, $A(0,-2a)$, $B(-a,-a)$, $D(d,0)$ y $E(d - a\cos\theta, -a \operatorname{sen}\theta)$, el baricentro de estos puntos, afectados con el mismo ”peso”, es el punto (con la notación 3-6):



$$G = \frac{1}{5}(A + B + C + D + E),$$

$$G\left(\frac{-a + 2d - a\cos\theta}{5}, \frac{-3a - a \operatorname{sen}\theta}{5}\right).$$

Para que el alambre esté en equilibrio, G ha de estar en la vertical de C, o sea $2d = a(1+\cos\theta)$. En particular, si $\theta = 30°$, $d = 3a/4$.
Applet-GeoGebra http://webpages.ull.es/users/amontes/geogebra/g1_09.html

Coordenadas baricéntricas

Con todo lo expuesto en esta sección, ya estamos en condiciones de definir las coordenadas baricéntricas.

3.20 Definición.- *Una **referencia baricéntrica** en un espacio afín $\mathcal{A}$ de dimensión n es un conjunto de $n+1$ puntos linealmente independientes $\Re = \{P_0, P_1, \ldots, P_n\}$. A los puntos que forman la referencia les denominamos **puntos básicos** de la referencia baricéntrica. Llamamos **coordenadas baricéntricas** de un punto $X \in \mathcal{A}$ a la única $(n+1)$–upla $(x^0, x^1, \ldots, x^n)$ de escalares tales que, para un punto $P \in \mathcal{A}$,*

$$\overrightarrow{PX} = x^0\overrightarrow{PP_0} + x^1\overrightarrow{PP_1} + \cdots + x^n\overrightarrow{PP_n}, \quad x^0 + x^1 + \cdots + x^n = 1_K.$$

Así, si $(x^0, x^1, \ldots, x^n)$ son las coordenadas baricéntricas del punto X, éste es el baricentro de los puntos $P_0, P_1, \ldots, P_n$ afectados de los ”pesos” (normalizados) $x^0, x^1, \ldots, x^n$.

La consistencia de esta definición, respecto a la unicidad de los escalares $(x^0, x^1, \ldots, x^n)$, así como de la independencia del punto P tomado, surgen de la independencia de los puntos $P_0, P_1, \ldots, P_n$.

Por notación, para indicar que $(x^0, x^1, \ldots, x^n)$ son las coordenadas baricéntricas de un punto X, escribiremos $X(x^0, x^1, \ldots, x^n)$.

Los puntos de la referencia baricéntrica tienen por coordenadas baricéntricas $P_0(1, 0, \ldots, 0)$, $P_1(0, 1, 0, \ldots, 0)$, ..., $P_n(0, \ldots, 0, 1)$.

Relación entre coordenadas cartesianas y baricéntricas

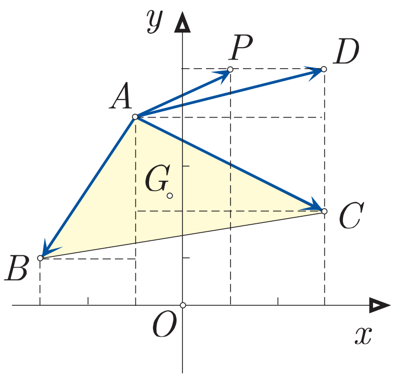
Si $(x^0, x^1, \ldots, x^n)$ son las coordenadas baricéntricas de un punto X respecto a una referencia baricéntrica $\Re = \{P_0, P_1, \ldots, P_n\}$, las coordenadas cartesianas de X en la referencia cartesiana $\mathcal{R} = \{P_0; \overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_n}\}$ son $(x^1, \ldots, x^n)$.

Recíprocamente, si $(x^1, \ldots, x^n)$ son las coordenadas cartesianas en la referencia cartesiana $\mathcal{R}$, las coordenadas baricéntricas en la referencia $\Re$ son $\big(1 - (x^1 + \cdots + x^n), x^1, \ldots, x^n\big)$.

3.21 Ejemplo.- *En plano afín real, respecto a la referencia cartesiana canónica, las coordenadas baricéntricas del punto $D(3, 5)$ respecto a la referencia $\Re = \{A(-1, 4), B(-3, 1), C(3, 2)\}$ (respecto al triángulo $\widehat{ABC}$) son $(9/8, -3/4, 5/8)$. El baricentro G del triángulo $\widehat{ABC}$ tiene coordenadas, respecto a la referencia canónica, $(-1/3, 7/3)$.*

$$\overrightarrow{AD} = \beta\overrightarrow{AB} + \gamma\overrightarrow{AC} \Rightarrow$$
$$(4, 1) = \beta(-2, -3) + \gamma(4, -2) \Rightarrow$$
$$\beta = -3/4, \gamma = 5/8 \Rightarrow$$
$$\alpha = 1 - \beta - \gamma = 9/8.$$

$$G = \frac{1}{3}(A + B + C) =$$
$$\frac{1}{3}\big((-1, 4) + (-3, 1) + (3, 2)\big) = \left(-\frac{1}{3}, \frac{7}{3}\right)$$



Applet-Cabri http://webpages.ull.es/users/amontes/cabri/g1-11.htm

3.22 Ejemplo.- *Con los datos del* Ejemplo 3.21*, el punto P cuyas coordenadas baricéntricas son $(5/4, -1/2, 1/4)$, tiene coordenadas en la referencia cartesiana canónica $(1, 5)$. Las coordenadas de P en la referencia cartesiana $\mathcal{R} = \{A; \overrightarrow{AB}, \overrightarrow{AC}\}$ son $(-1/2, 1/4)$.*

Antes de tratar unos teoremas clásicos, vamos a usar la noción de baricentro de un subconjunto de puntos para dar una caracterización de variedad lineal.

En Geometría Elemental se define un plano por la condición de que si contiene a dos puntos también contiene a la recta que los une. Como hemos adoptado ya otra definición de variedad lineal (y en particular de plano) se podría pensar que desde el punto de vista adoptado aquí (Definición 2.13) la definición elemental se va a convertir en una proposición, como vamos a ver a continuación.

3.23 Proposición.- *Un subconjunto A no vacío de un espacio afín $(\mathcal{A}, \mathbf{E})$ es una variedad lineal si y sólo si para toda familia de puntos ponderados $\{(P_k, \alpha^k)\}_{k=1}^r$ de A, con $\sum_{k=1}^r \alpha^k = 1_K$, el baricentro $\sum_{k=1}^r \alpha^k P_k$ pertenece a A.*

Demostración.- Si A es una variedad lineal y P es un punto de A, el baricentro de los puntos ponderados $\{(P_k, \alpha^k)\}_{k=1}^r$:

$$B = P + \bar{\alpha}^1 \overrightarrow{PP_1} + \cdots + \bar{\alpha}^r \overrightarrow{PP_r} \quad (\bar{\alpha}^k = \alpha^k/(\alpha^1 + \cdots + \alpha^r),$$

pertenece a A, ya que los vectores $\overrightarrow{PP_k}$ pertenecen al subespacio director de la variedad lineal A.

Recíprocamente, para de demostrar que A es una variedad lineal, veamos que si $P \in \mathsf{A}$, el conjunto $\mathbf{A} = \{\overrightarrow{PX} \in \mathbf{E} / X \in \mathsf{A}\}$ es un subespacio de $\mathbf{E}$.

Sean $P, X, Y \in \mathsf{A}$ y probemos que, para todo $\lambda, \mu \in K$, $\lambda\overrightarrow{PX} + \mu\overrightarrow{PY} \in \mathbf{A}$
Sea el punto Z tal que

$$\overrightarrow{PZ} = \lambda\overrightarrow{PX} + \mu\overrightarrow{PY} = \lambda\overrightarrow{PX} + \mu\overrightarrow{PY} + (1 - (\lambda + \mu))\overrightarrow{PP},$$

entonces, Z es el baricentro de los puntos X, Y, P afectados de los "pesos" λ, μ y $1 - (\lambda + \mu)$, respectivamente. Por tanto, $Z \in \mathsf{A}$, es decir $\overrightarrow{PZ} \in \mathbf{A}$. ◪

Ya que el baricentro de m puntos ponderados puede ser obtenido reiterando la determinación de un par de puntos ponderados, podemos afirmar:

3.24 Proposición.- *Un subconjunto A, que contiene más de un punto, de un espacio afín $\mathcal{A}$ es una variedad lineal si sólo si para todo par de puntos distintos $P, Q \in \mathsf{A}$, el conjunto A contiene a la recta determinada por P y Q, esto es al conjunto de puntos de la forma $tP + (1-t)Q$, $t \in K$.*

◪

Razón simple

3.25 Definición.- *Se llama **razón simple** de tres puntos P_1, P_2 y P_3 ($P_1 \neq P_2$) alineados en un espacio afín al único escalar, que denotaremos por $(P_1\ P_2\ P_3)$, tal que*

$$\overrightarrow{P_1P_3} = (P_1\ P_2\ P_3)\ \overrightarrow{P_1P_2}.$$

Consecuencia inmediata de la Definición de baricentro, pág. 40 es que P_1 es el baricentro de P_2 y P_3 afectados de los "pesos" $-(P_1\ P_2\ P_3)$ y 1, respectivamente.

Podemos expresar la razón simple en términos de las coordenadas cartesianas de los puntos una recta: si los tres puntos P_1, P_2 y P_3 ($P_1 \neq P_2$) están en una recta $\mathcal{L}$ que pasa por el punto P y tiene vector director $\vec{u}$, es decir, $\{P; \vec{u}\}$

es una referencia cartesiana sobre $\mathcal{L}$, entonces $\overrightarrow{PP_i} = \lambda_i \vec{\boldsymbol{u}}$, $i = 1, 2, 3$ ($\lambda_1 \neq \lambda_2$),

$$\overrightarrow{P_1P_2} = (\lambda_2 - \lambda_1)\vec{\boldsymbol{u}}, \quad \overrightarrow{P_1P_3} = (\lambda_3 - \lambda_1)\vec{\boldsymbol{u}}, \quad \overrightarrow{P_1P_3} = \frac{\lambda_3 - \lambda_1}{\lambda_2 - \lambda_1}\,\overrightarrow{P_1P_2}.$$

Luego, la razón simple de los tres puntos P_1, P_2 y P_3 respecto a la referencia $\{P; \vec{\boldsymbol{u}}\}$ es:

$$(P_1\ P_2\ P_3) = \frac{\lambda_3 - \lambda_1}{\lambda_2 - \lambda_1}.$$

Si tomamos $\{P_1; \overrightarrow{P_1P_2}\}$ como referencia, se tiene que $(P_1\ P_2\ P_3) = \lambda_3$.

El baricentro de los puntos P y Q afectados de los "pesos" α y β es el punto G tal que:

$$(G\ P\ Q) = -\frac{\alpha}{\beta}.$$

Esto surge de $\alpha\overrightarrow{GP} + \beta\overrightarrow{GQ} = \vec{\boldsymbol{0}} \Rightarrow \overrightarrow{GQ} = -\frac{\alpha}{\beta}\overrightarrow{GP} = (G\ P\ Q)\ \overrightarrow{GP}$.

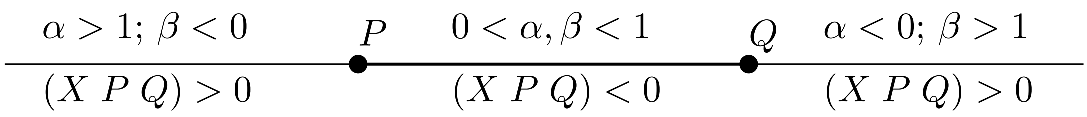
La razón simple de tres puntos distintos P, Q y R depende del orden en que se elijan. Como hay seis permutaciones de tres puntos distintos, hay seis formas en que la razón simple de tres puntos distintos puede escribirse. Los posibles seis valores, si $(P\ Q\ R) = \rho$, son:

$$(P\ Q\ R) = \rho, \qquad (P\ R\ Q) = \frac{1}{\rho}, \qquad (Q\ P\ R) = 1 + \rho,$$

$$(Q\ R\ P) = \frac{1}{1+\rho}, \quad (R\ Q\ P) = \frac{\rho}{1+\rho}, \quad (R\ P\ Q) = \frac{1+\rho}{\rho}.$$

3.26 Nota.- Los valores distintos de la razón simple obtenidos al permutar los tres puntos ponen de manifiesto que la definición 3.25 de razón simple es una de las seis posibles que se pueden dar, es por lo que el la literatura cada autor puede tomar una distinta. Aunque, para cualquiera de las que se tome, el resto se pueden obtener a partir de ella.

En el espacio afín real, las coordenadas baricéntricas en un recta nos permiten conocer la posición de un punto X respecto a otros dos P y Q. Si (α, β) son las coordenadas baricéntricas de X en la recta PQ ($\overrightarrow{PX} = \beta\overrightarrow{PQ}$, $\alpha + \beta = 1$), se tiene los signos de sus coordenadas y de la razón simple $(X\ P\ Q)$, como se indican en la figura:



Esto nos permite definir un ***segmento*** de extremos P y Q en una recta real, que denotamos por $\overline{PQ}$, como el conjunto:

$$\overline{PQ} = \{X \in \mathcal{A}/X = \alpha P + \beta Q,\ \alpha + \beta = 1,\ 0 \leq \alpha, \beta \leq 1\}.$$

Los dos resultados siguientes se conocen como teorema de Tales (Tales de Mileto, 624-547 a.C.):

1. Cuando dos rectas paralelas cortan a dos rectas secantes, determinan en éstas segmentos proporcionales.
2. El ángulo inscrito en un semicircunferencia es recto.

Vamos a establecer el primero en el caso más general de un espacio afín de dimensión n, usando la ecuación cartesiana de un hiperplano (se da otra demostración, usando proyecciones, en la Proposición 4.33).

3.27 Proposición [Teorema de Tales].- *Sean $\mathcal{H}_1, \mathcal{H}_2$ y $\mathcal{H}_3$ tres hiperplanos paralelos de un espacio afín $(\mathcal{A}, \boldsymbol{E}, +)$ de dimensión n. Una recta $\mathcal{L}$ no paralela a los hiperplanos, los corta en los puntos P_1, P_2 y P_3, respectivamente, entonces la razón simple $(P_1\ P_2\ P_3)$ sólo depende de los hiperplanos y no de la recta que los contiene.*

Demostración.- Sea $\boldsymbol{H} \dot{\subset} \boldsymbol{E}$ el subespacio director de los tres planos paralelos y $\{\vec{\boldsymbol{u}}_1, \dots, \vec{\boldsymbol{u}}_{n-1}\} \subset \boldsymbol{F}$, $n-1$ vectores independientes.

Puntos genéricos de los hiperplanos se pueden expresar en la forma:

$$X = Q_i + \lambda_i^1 \vec{\boldsymbol{u}}_1 + \cdots + \lambda_i^{n-1} \vec{\boldsymbol{u}}_{n-1} \quad (i = 1, 2, 3)$$

Es decir, para una referencia cartesiana $\mathcal{R} = \{O; \vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$, respecto a la cual $Q_i(q_i^1, \dots, q_i^n)$ y $\vec{\boldsymbol{u}}_j = \sum_{k=1}^{n} a_j^k \vec{\boldsymbol{e}}_k$ $(i = 1, 2, 3; j = 1, \dots, n-1)$, las ecuaciones paramétricas de los hiperplanos son:

$$\begin{array}{l} x^1 - q_i^1 = \lambda_i^1 a_1^1 + \cdots + \lambda_i^{n-1} a_{n-1}^1 \\ x^2 - q_i^2 = \lambda_i^1 a_1^2 + \cdots + \lambda_i^{n-1} a_{n-1}^2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x^n - q_i^n = \lambda_i^1 a_1^n + \cdots + \lambda_i^{n-1} a_{n-1}^n \end{array}$$

Como a cada punto del hiperplano $\mathcal{H}_i$ de coordenadas $(x^1, \dots, x^n)$ le corresponde unos únicos escalares $(\lambda_i^1, \dots, \lambda_i^{n-1})$, debe verificarse que:

$$\begin{vmatrix} x^1 - q_i^1 & a_1^1 & \cdots & a_{n-1}^1 \\ x^2 - q_i^2 & a_1^2 & \cdots & a_{n-1}^2 \\ \cdots & \cdots & \cdots & \cdots \\ x^n - q_i^n & a_1^n & \cdots & a_{n-1}^n \end{vmatrix} = 0.$$

Esta es la ecuación cartesiana del hiperplano $\mathcal{H}_i$. Por tanto, desarrollando este determinante, las ecuaciones de los tres hiperplanos son:

$$\begin{array}{ll} \mathcal{H}_1 : & \alpha_1 x^1 + \cdots + \alpha_n x^n = \beta_1 \\ \mathcal{H}_2 : & \alpha_1 x^1 + \cdots + \alpha_n x^n = \beta_2 \\ \mathcal{H}_3 : & \alpha_1 x^1 + \cdots + \alpha_n x^n = \beta_3 \end{array} \tag{3-7}$$

Vamos a demostrar que la razón simple $(P_1\ P_2\ P_3)$, de los puntos $P_i(p_i^1, \dots, p_i^n)$ de intersección de una recta $\mathcal{L}$ con los hiperplanos, sólo depende de los escalares β_1, β_2 y β_3.

Tomemos sobre la recta $\mathcal{L}$ la referencia cartesiana $\{P_1; \overrightarrow{P_1P_2}\}$, respecto a la cual las coordenadas cartesianas de los puntos P_1, P_2 y P_3 son $0, 1$ y ρ, respectivamente, siendo ρ el escalar tal que $\overrightarrow{P_1P_3} = \rho\, \overrightarrow{P_1P_2}$, o sea $\rho = (P_1\ P_2\ P_3)$.

Al pertenecer $P_i \in \mathcal{H}_i$ $(i = 1, 2, 3)$ se tiene que:

$$\alpha_1 p_1^1 + \cdots + \alpha_n p_1^n = \beta_1$$
$$\alpha_1 p_2^1 + \cdots + \alpha_n p_2^n = \beta_2$$

$$\alpha_1(p_1^1 + \rho(p_2^1 - p_1^1)) + \cdots + \alpha_n(p_1^n + \rho(p_2^n - p_1^n)) = \beta_3.$$

De la tercera relación, usando las dos primeras, se obtiene que:

$$\beta_1 + \rho\beta_2 - \rho\beta_1 = \beta_3 \Rightarrow (P_1\ P_2\ P_3) = \rho = \frac{\beta_3 - \beta_1}{\beta_2 - \beta_1}.$$

3.28 Nota.- Las ecuaciones cartesianas (3-7) de hiperplanos paralelos nos indican que para que dos hiperplanos:

$$\alpha_1 x^1 + \cdots + \alpha_n x^n = \gamma_1, \qquad \beta_1 x^1 + \cdots + \beta_n x^n = \gamma_2,$$

sean paralelos se debe verificar que $\alpha_1/\beta_1 = \cdots = \alpha_n/\beta_n$.

En espacio euclídeos y respecto a una referencia rectangular, si $a_1x^1 + \cdots + a_nx^n = b$ es la ecuación de hiperplano el vector de componentes $(a_1, \ldots, a_n)$ es perpendicular al hiperplano (Proposición 6.6).

Otros teoremas clásicos, en cuya demostración se hace uso del valor de la razón simple de tres puntos en términos de las coordenadas baricéntricas de uno de ellos respecto a los otros dos, son los Menelao y Ceva.

3.29 Proposición[Teorema de Menelao].- *Sean A, B, C tres puntos linealmente independientes, y sean D, E y F puntos de las rectas determinadas por B y C, C y A, A y B, respectivamente, que están alineados y sean diferentes de A, B y C. Entonces:*

$$(D\,B\,C)\,(E\,C\,A)\,(F\,A\,B) = 1.$$

Demostración.- Utilizaremos coordenadas baricéntricas respecto a la referencia $\{A, B, C\}$. Las coordenadas de los puntos D, E y F son:

$$D(0, \alpha_2, \alpha_3), \qquad E(\beta_1, 0, \beta_3), \qquad F(\gamma_1, \gamma_2, 0).$$

Entonces,

$$(D\,B\,C) = -\frac{\alpha_2}{\alpha_3}, \quad (E\,C\,A) = -\frac{\beta_2}{\beta_1}, \quad (F\,A\,B) = -\frac{\gamma_1}{\gamma_2}.$$

Luego,

$$(D\,B\,C)\,(E\,C\,A)\,(F\,A\,B) = -\frac{\alpha_2\beta_2\gamma_1}{\alpha_3\beta_1\gamma_2}.$$

Por otra parte, los puntos D, E y F estén alineados si

$$\begin{vmatrix} 0 & \alpha_2 & \alpha_3 \\ \beta_1 & 0 & \beta_3 \\ \gamma_1 & \gamma_2 & 0 \end{vmatrix} = \alpha_2\beta_2\gamma_1 + \alpha_3\beta_1\gamma_2 = 0 \Leftrightarrow -\frac{\alpha_2\beta_2\gamma_1}{\alpha_3\beta_1\gamma_2} = 1.$$

La condición de alineamiento de los puntos D, E y F en el Teorema de Menelao la podemos escribir de la forma (suprimiendo la raya encima para expresar el segmento de extremos dos puntos):

$$\frac{BD}{DC}\,\frac{CE}{EA}\,\frac{AF}{FB} = -1.$$

Esto es así si suponemos definida la distancia entre puntos, para poder dar la longitud de un segmento, lo cual sería lícito en el plano euclídeo.

3.30 Proposición[Teorema de Ceva].- *Sean A, B, C puntos linealmente independientes en un espacio afín $\mathcal{A}$ de dimensión 2. Dado $P \in \mathcal{A}$, $P \notin \{A, B, C\}$, sean $D = BC \cap AP$, $E = CA \cap BP$ y $F = AB \cap CP$. Entonces:*

$$(D\,B\,C)\,(E\,C\,A)\,(F\,A\,B) = -1.$$

Demostración.- Sea $P(\alpha, \beta, \gamma)$ las coordenadas baricéntricas de P, respecto a $\{A, B, C\}$, entonces

$$D\left(0, \frac{\beta}{1-\alpha}, \frac{\gamma}{1-\alpha}\right) \quad \Rightarrow \quad (D\,B\,C) = -\frac{\beta}{\gamma},$$
$$E\left(\frac{\alpha}{1-\beta}, 0, \frac{\gamma}{1-\beta}\right) \quad \Rightarrow \quad (E\,C\,A) = -\frac{\gamma}{\alpha},$$
$$F\left(\frac{\alpha}{1-\gamma}, \frac{\beta}{1-\gamma}, 0\right) \quad \Rightarrow \quad (F\,A\,B) = -\frac{\alpha}{\beta}.$$

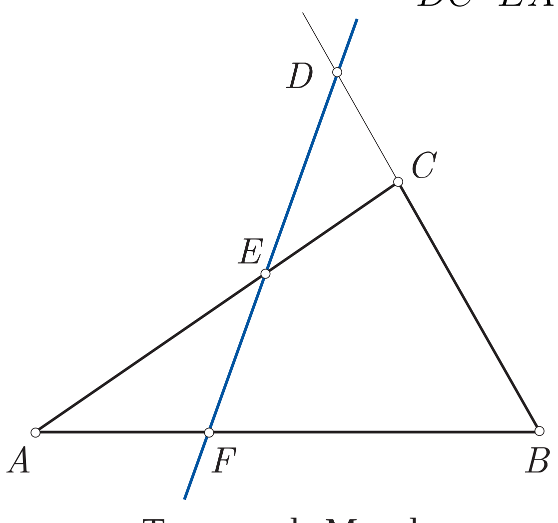
Con lo que

$$(D\,B\,C)\,(E\,C\,A)\,(F\,A\,B) = -\frac{\beta\gamma\alpha}{\gamma\alpha\beta} = -1.$$

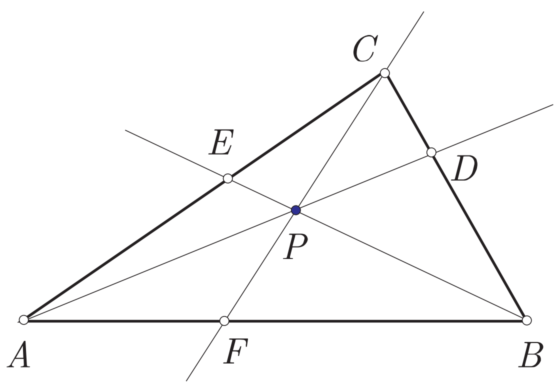
■

La condición de concurrencia de las rectas AD, BE y CF en el Teorema de Ceva la podemos escribir de la forma, si suponemos definida la distancia entre puntos para poder dar la longitud de un segmento, para lo cual necesitamos estar en el plano euclídeo:

$$\frac{BD}{DC}\,\frac{CE}{EA}\,\frac{AF}{FB} = 1.$$



Teorema de Menelao



Teorema de Ceva

Vamos a establecer el conocido com Teorema del "hexagrama místico" de Pascal para un hexágono inscrito en una circunferencia, utilizando el Teorema de Menelao. Es importante hacer notar que no podemos de hablar de circunferencias hasta que sea introducido el concepto de plano euclídeo

3.31 Proposición[Teorema de Pascal].- *Los puntos de intersección de los lados opuestos de un hexágono inscrito en una circunferencia están alineados.*

Demostración.- [2]

Dada una circunferencia en el plano, sea $ABCDEF$ un hexágono inscrito (los seis vértices pueden estar dispuestos en cualquier orden) y sean los puntos $P = AB \cap DE, Q = BC \cap EF, R = CD \cap FA$. Tenemos que probar que P, Q y R están alineados.

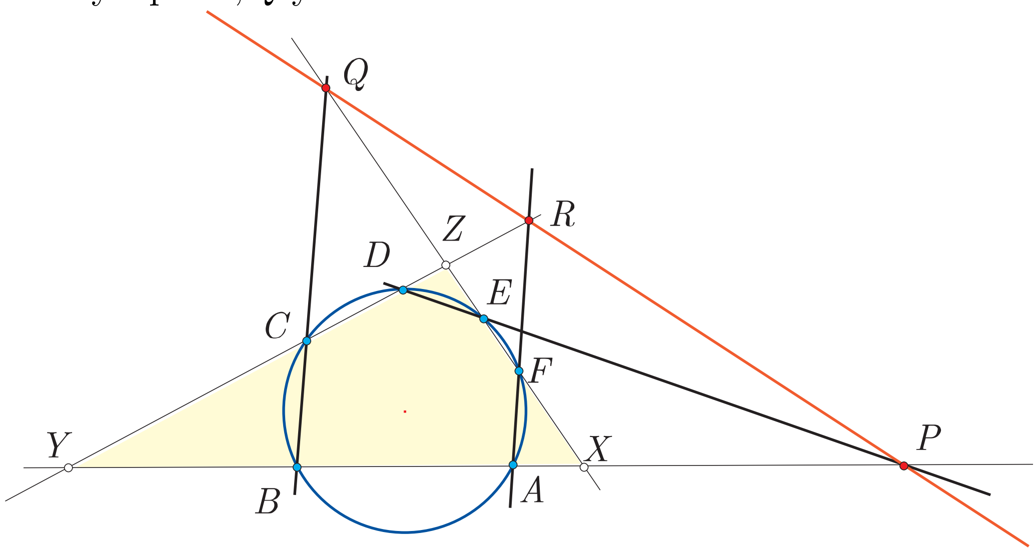
Los tres lados AB, CD y EF (el primero, el tercero y el quinto), forman un triángulo XYZ, y los tres lados BC, DE y FA (el segundo, el cuarto y el sexto) son rectas transversales al triángulo XYZ y aplicando el Teorema de Menelao, se tiene:

$$\frac{ZQ}{XQ}\cdot\frac{XB}{YB}\cdot\frac{YC}{ZC} = 1, \qquad \frac{XP}{YP}\cdot\frac{YD}{ZD}\cdot\frac{ZE}{XE} = 1, \qquad \frac{YR}{ZR}\cdot\frac{ZF}{XF}\cdot\frac{XA}{YA} = 1.$$

Multiplicando miembro a miembro estas igualdades y simplificando, utilizando el valor de la potencia de un punto respecto a una circunferencia para obtener las relaciones $XA \cdot XB = XE \cdot XF$, $YC \cdot YD = YA \cdot YB$, $ZE \cdot ZF = ZC \cdot ZD$, se obtiene que:

$$\frac{ZQ}{XQ}\cdot\frac{XP}{YP}\cdot\frac{YR}{ZR} = 1.$$

Se concluye que P, Q y R están alineados. ◾



Coordenadas baricéntricas homogéneas

Que $(x^0, x^1, \ldots, x^n)$ sean las coordenadas baricéntricas de un punto $X \in \mathcal{A}$, respecto a una referencia baricéntrica $\Re = \{A_0, A_1, \ldots, A_n\}$, significa que X es el baricentro de los puntos $A_0, A_1, \ldots, A_n$ afectados de los "pesos" $x^0, x^1, \ldots, x^n$ $(x^0 + x^1 + \cdots + x^n = 1_K)$; pero como X sigue siendo el baricentro de los

[2] Existen otras demostraciones de este teorema en Tan, K..- "Various proofs of Pascal's Theorem", Math. Mag. 38, 22-30 (1965) (http://www.jstor.org/stable/2688012).

puntos $A_0, A_1, \ldots, A_n$ y se multiplican sus "pesos" asociados por un mismo escalar $\lambda \neq 0_K$, se pueden considerar cualquier n–upla proporcional a la $(x^0, x^1, \ldots, x^n)$ para caracterizar al punto X. A tal elección denominaremos ***coordenadas baricéntrica homogéneas*** del punto X respecto a la referencia baricéntrica $\Re$, que denotaremos por $(x^0 : x^1 : \cdots : x^n)$.

Dadas las coordenadas baricéntricas $(x^0 : x^1 : \cdots : x^n)$ de un punto, obtenemos sus coordenadas baricéntricas (absolutas) sin mas que dividir todas sus componentes por $x^0 + x^1 + \cdots + x^n$.

Ecuaciones baricéntricas de una variedad lineal

Tomemos un punto $X \in \mathcal{F}$ ($\mathcal{F}$ variedad lineal de $\mathcal{A}$) de coordenadas baricéntricas homogéneas $(x^0 : x^1 : \cdots : x^n)$, respecto a una referencia baricéntrica dada en $\mathcal{A}$, $\{A_0, A_1, \ldots, A_n\}$. Si $\{P_0, P_1, \ldots, P_r\}$ son puntos independientes en $\mathcal{F}$, cuyas coordenadas baricéntricas homogéneas son $(a_j^0 : a_j^1 : \cdots : a_j^n)$ $(j = 0, 1, \ldots, r)$, se tiene

$$\left.\begin{array}{rcl} X & = & \sum\limits_{j=0}^{r} \lambda^j P_j = \sum\limits_{j=0}^{r}\sum\limits_{i=0}^{n} \lambda^j a_j^i A_i \\ X & = & \sum\limits_{i=0}^{n} x^i A_i \end{array}\right\} \Rightarrow \rho x^i = \sum_{j=0}^{r} \lambda^j a_j^i \quad (i = 0, 1, \ldots n).$$

Para que este sistema de ecuaciones homogéneo de $n+1$ ecuaciones y $r+2$ incógnitas $(\rho, \lambda^0, \lambda^1, \ldots, \lambda^n)$, tenga solución, elegido un menor de orden $r+1$ distinto de cero (que se está garantizado por ser los puntos $P_0, P_1, \ldots, P_r$ linealmente independientes y la matriz (a_j^i) tiene rango $r+1$), todos los menores de orden $r+2$, obtenidos orlándole la columna restante y cada una de las filas restantes, han de ser nulos. Se obtienen así $n-r$ ecuaciones en las variables $x^0, x^1, \ldots, x^n$, denominadas ***ecuaciones baricéntricas*** de la variedad lineal $\mathcal{F}$:

$$\begin{array}{ccccccccc} \alpha_0^1 x^0 & + & \alpha_1^1 x^1 & + & \cdots & + & \alpha_n^1 x^n & = & 0 \\ \alpha_0^2 x^0 & + & \alpha_1^2 x^1 & + & \cdots & + & \alpha_n^2 x^n & = & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_0^{n-r} x^0 & + & \alpha_1^{n-r} x^1 & + & \cdots & + & \alpha_n^{n-r} x^n & = & 0 \end{array}$$

La ecuación baricéntrica de un hiperplano $\mathcal{H}$, variedad lineal de dimension $n-1$ en un espacio afín $\mathcal{A}$ de dimension n, es:

$$\alpha_0 x^0 + \alpha_1 x^1 + \cdots + \alpha_n x^n = 0.$$

En particular, el hiperplano generado por n puntos de un sistema de coordenadas baricéntricas $\{A_0, A_1, \ldots, A_n\}$, por ejemplo, el generados por todos los puntos menos el A_k, tiene de ecuación baricéntrica:

$$x^k = 0.$$

En el plano afín y en espacio afines tridimensional, los hiperplanos son las rectas y los planos, que tiene por ecuaciones, respectivamente:

$$ax + by + cz = 0, \qquad ax + by + cz + dt = 0,$$

si tomamos como coordenadas baricéntricas homogéneas de un punto en ambos $(x : y : z)$ y $(x : y : z : t)$.

Puntos del infinito del espacio afín real en coordenadas baricéntricas homogéneas

Sean $(\mathcal{A}, \mathbb{R}^n, +)$ un espacio afín real de dimensión n y $\Re = \{A_0, A_1, \dots, A_n\}$ una referencia baricéntrica: todo punto $X \in \mathcal{A}$ tiene coordenadas baricéntricas homogéneas $(x^0 : x^1 : \cdots : x^n)$, respecto a la referencia $\Re$. Consideremos la recta $\mathfrak{L}$ que pasa por $P(p^0 : p^1 : \cdots : p^n)$ y $Q(q^0 : q^1 : \cdots : q^n)$, cuyas coordenadas las podemos elegir de tal forma que se verifique que

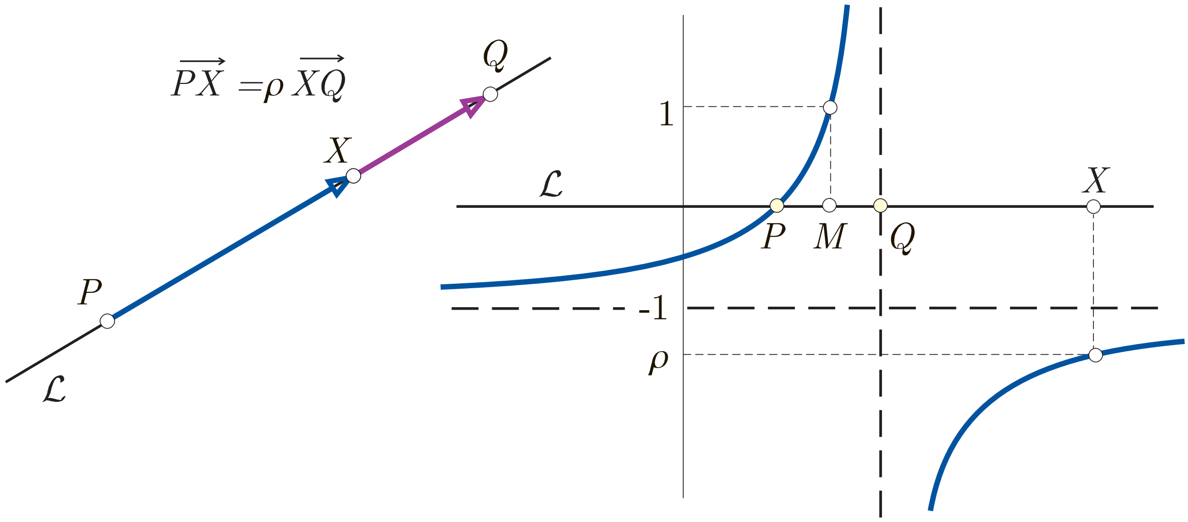
$$p^0 + p^1 + \cdots + p^n = q^0 + q^1 + \cdots + q^n.$$

Sobre la recta $\mathfrak{L}$, el punto X que separa P y Q en la razón ρ (es decir [3], $\overrightarrow{PX} = \rho\overrightarrow{XQ}$ ó $\overrightarrow{XP} + \rho\overrightarrow{XQ} = \vec{\boldsymbol{0}}$) tiene coordenadas baricéntricas:

$$(p^0 + \rho q^0 : p^1 + \rho q^1 : \cdots : p^n + \rho q^n), \tag{3-8}$$

pues de $\overrightarrow{XP} = p^0\overrightarrow{XA_0} + \cdots + p^n\overrightarrow{XA_n}$ y $\overrightarrow{XQ} = q^0\overrightarrow{XA_0} + \cdots + q^n\overrightarrow{XA_n}$ se sigue que:

$$(p^0 + \rho q^0)\overrightarrow{XA_0} + \cdots + (p^n + \rho q^n)\overrightarrow{XA_n} = \vec{\boldsymbol{0}}.$$



Para valores particulares de ρ, se obtienen los siguientes puntos de la recta $\mathfrak{L}$:

$\rho = 0, X = P$; $\rho = 1, X$ es el punto medio M de P y Q; $\rho \to \infty, X \to Q$;
$\rho = -1, X$ tiende al punto del infinito de la recta $\mathfrak{L}$.

De este último caso se obtiene que las coordenadas baricéntricas del punto del infinito de la recta determinada por los puntos P y Q son:

$$(p^0 - q^0 : p^1 - q^1 : \cdots : p^n - q^n),$$

cuya suma es cero, por lo que dicho punto no tiene coordenadas baricéntricas absolutas. Como $P \neq Q$, un punto del infinito no puede tener coordenadas de la forma $(0 : \cdots : 0)$.

El punto del infinito de la recta A_iA_j, determinada por dos puntos básicos de la referencia $\Re$, tiene coordenadas baricéntricas homogéneas:

$$(0 : \cdots : \overset{i)}{1} : \cdots : \overset{j)}{-1} : \cdots : 0) = (0 : \cdots : \overset{i)}{-1} : \cdots : \overset{j)}{1} : \cdots : 0).$$

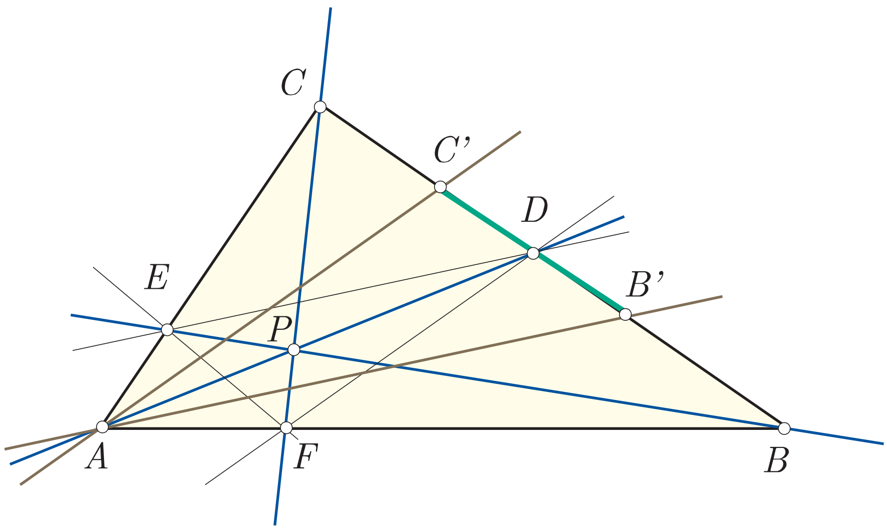
Las coordenadas baricéntricas homogéneas $(x^0 : x^1 : \cdots : x^n)$, respecto a una referencia $\Re$, de los puntos del infinito de las rectas en el espacio afín real

(3) Que el punto X divida a los puntos P y Q en la razón ρ, quiere decir que X es el baricentro de P y Q afectados de los "pesos" 1 y ρ; o bien, la razón simple $(X\ P\ Q) = -\rho$.

satisfacen a la ecuación:
$$x^0 + x^1 + \cdots + x^n = 0,$$
se dice que es la ecuación del ***hiperplano del infinito***, respecto a la referencia $\Re$.

Veamos un ejemplo del uso de coordenadas homogéneas en el plano y de los puntos del infinito.

3.32 Ejemplo.- *Dado un triángulo $\widehat{ABC}$ y un punto P en el plano afín real, sean D, E y F los pies de las cevianas de P, esto es, los puntos en los que las rectas AP, BP y CP cortan a las rectas BC, CA y AB, respectivamente.*
La paralela por A a DE corta a BC en B' y la paralela por A a DC corta a BC en C'; entonces D es el punto medio del segmento B'C'.



Una verificación de este hecho usando los teoremas de Ceva y Tales, puede ser la siguiente:

Como las rectas AD, BE y CF son concurrentes (en P) por el Teorema de Ceva se tiene que:
$$\frac{BD}{DC}\,\frac{CE}{EA}\,\frac{AF}{FB} = -1.$$
Que podemos reescribir de la forma:
$$\frac{DB\;AF}{FB} = \frac{CD\;EA}{CE}.$$

Por el Teorema de Tales, al ser $DF \| B\text{'}A$, se tiene que $B\text{'}D = \dfrac{DB\;AF}{FB}$; y por ser $DE \| C\text{'}A$, se tiene que $DC\text{'} = \dfrac{CD\;EA}{CE}$. Por tanto la condición de Ceva es equivalente a $B\text{'}D = DC\text{'}$, es decir, D es el punto medio de B' y C'.

Ahora utilizamos coordenadas baricéntrica homogéneas, respecto a la referencia dada por los vértices del triángulo $\widehat{ABC}$, que denotamos por $(x : y : z)$.

Las ecuaciones de los lados del triángulo son $BC \equiv x = 0$, $CA \equiv y = 0$ y $AB \equiv z = 0$. Si $P(\alpha : \beta : \gamma)$ entonces $D(0 : \beta : \gamma)$, $E(\alpha : 0 : \gamma)$, y $F(\alpha : \beta : 0)$.

La recta por A paralela a DE tiene el mismo punto del infinito que el de ésta, que es el punto $(\alpha(\beta+\gamma) : -\beta(\alpha+\gamma) : \gamma(-\alpha+\beta))$, solución de las ecuaciones:

$$\begin{vmatrix} x & y & z \\ 0 & \beta & \gamma \\ \alpha & 0 & \gamma \end{vmatrix} = \beta\gamma x + \alpha\gamma y - \alpha\beta z = 0, \qquad x+y+z=0.$$

La recta paralela por A a DE es $\gamma(\alpha-\beta)y - \beta(\alpha+\gamma)z = 0$. Y el punto de intersección con BC es B'$(0 : \beta(\alpha+\gamma) : \gamma(\alpha-\beta))$.

Cálculo similares nos dan que el punto de intersección con BC de la resta paralela por A a DF es: $C'(0 : \beta(\alpha-\gamma) : (\alpha+\beta)\gamma)$.

El punto medio de los puntos B' y C', se determina por la fórmula (3-8), para $\rho = 1$, que es:

$$(\beta(\alpha-\gamma)+\gamma(\alpha+\beta))\Big(0 : \beta(\alpha+\gamma) : \gamma(\alpha-\beta)\Big) + (\gamma(\alpha-\beta)+\beta(\alpha+\gamma))\Big(0 : \beta(\alpha-\gamma) : \gamma(\alpha+\beta)\Big)$$

$$= \left(0 : -2\alpha^2\beta(\beta+\gamma) : -2\alpha^2\gamma(\beta+\gamma)\right).$$

Que en virtud de la proporcionalidad de las coordenadas homogéneas es $(0 : \beta : \gamma)$; es decir, el punto D.

Los pies de las cevianas de un punto $P(\alpha : \beta : \gamma)$ dado en coordenadas baricéntricas respecto a un triángulo ABC en el plano, es decir, los puntos de intersección de cada recta que pasa por P y un vértice con el lado opuesto a tal vértice son:

$$D(0 : \beta : \gamma), \qquad E(\alpha : 0 : \gamma), \qquad F(\alpha : \beta : 0).$$

Recíprocamente, si éstas son, respectivamente, las coordenadas de tres puntos D, E, F, entonces las rectas AD, BE y CF se cortan en el punto de coordenadas $(\alpha : \beta : \gamma)$.

Utilizamos esto para determinar que las alturas de un triángulo son concurrentes y, de paso, obtenemos sus coordenadas baricéntricas homogéneas, respecto al triángulo en cuestión.

3.33 Ejemplo.- *Si por a, b y c se denotan las longitudes de los lados $\overline{BC}, \overline{CA}$ y $\overline{AB}$ de un triángulo ABC, las alturas del mismo concurren en un punto, llamado ortocentro, de coordenadas baricéntricas homogéneas*

$$(\operatorname{tag}\hat{A} : \operatorname{tag}\hat{B} : \operatorname{tag}\hat{C}),$$

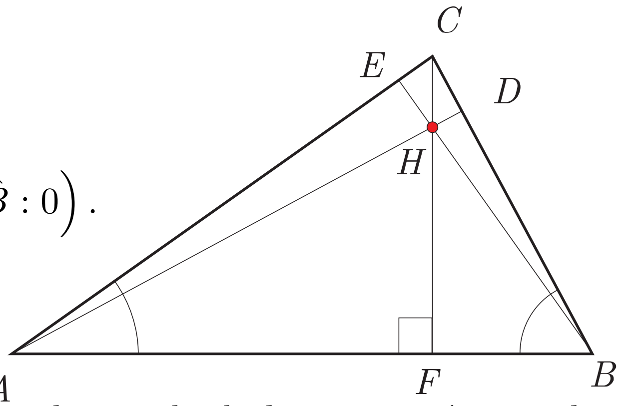
siendo $\hat{A}, \hat{B}$ y $\hat{C}$ los ángulos en los vértices A, B y C, respectivamente.

Sea F el pie de la altura desde el vértice C, entonces las proyecciones de los lados AC y BC sobre la recta AB son, respectivamente, $AF = b\cos\hat{A}$ y $FB = a\cos\hat{B}$; por lo que, las coordenadas de F son de la forma $(a\cos\hat{B} : b\cos\hat{A} : 0)$, que podemos poner, utilizando el teorema del seno:

$$\frac{a}{\operatorname{sen}\hat{A}} = \frac{b}{\operatorname{sen}\hat{B}} = \frac{c}{\operatorname{sen}\hat{C}},$$

y dividiendo todas los componentes por el factor $\cos\hat{A}\cos\hat{B}$ de cualquiera de las formas siguientes:

$$\left(\frac{\operatorname{sen}\hat{A}}{\cos\hat{A}} : \frac{\operatorname{sen}\hat{B}}{\cos\hat{B}} : 0\right) = \left(\operatorname{tag}\hat{A} : \operatorname{tag}\hat{B} : 0\right).$$



Procediendo de forma similar con las alturas desde los otros vértices, los pies de las tres alturas son:

$D\big(0 : \operatorname{tag}\hat{B} : \operatorname{tag}\hat{C}\big)$, $\quad E\big(\operatorname{tag}\hat{A} : 0 : \operatorname{tag}\hat{C}\big)$, $\quad F\big(\operatorname{tag}\hat{A} : \operatorname{tag}\hat{B} : 0\big)$.

Con lo que las tres alturas concurren el punto denominado ortocentro:

$$\big(\operatorname{tag}\hat{A} : \operatorname{tag}\hat{B} : \operatorname{tag}\hat{C}\big).$$

3.34 Ejemplo.- *Dado un triángulo $\widehat{ABC}$, construir un triángulo $\widehat{DEF}$ tal que el punto medio de A y E sea F, el punto medio de B y F sea D y el punto medio D y C sea E.*

En la referencia baricéntrica $\Re = \{A, B, C\}$, sean las coordenadas baricéntricas homogéneas de $E(u : v : w)$. Para determinar las coordenadas del punto F, punto medio de A y E, normalicemos las coordenadas de estos puntos, de tal forma que las sumas de sus componentes sean iguales: $A(u + v + w : 0 : 0)$ y $E(u : v : w)$; luego, $F = A + E = (2u + v + w : v : w)$

El punto medio D de B y F es:

$(0 : 2u+2v+2w : 0)+(2u+v+w : v : w) = (2u+v+w : 2u+3v+2w : w)$.

Finalmente, las coordenadas del punto medio D de $C(0 : 0 : 4u + 4v + 4w)$ y F son $(2u + v + w : 2u + 3v + 2w : 4u + 4v + 5w)$, que deben coincidir con las de $E(u : v : w)$, salvo un factor de proporcionalidad.

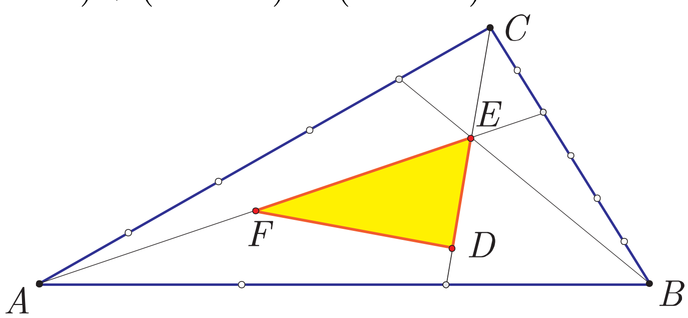
Tenemos que resolver el siguiente sistema homogéneo de ecuaciones:

$(2 - \lambda)u + v + w = 0, \quad 2u + (3 - \lambda)v + 2w = 0, \quad 4u + 4v + (5 - \lambda)w = 0.$

Para que tenga solución, distinta de la trivial $u = v = w = 0$, se tiene que verificar que:

$$\begin{vmatrix} 2-\lambda & 1 & 1 \\ 2 & 3-\lambda & 2 \\ 4 & 4 & 5-\lambda \end{vmatrix} = -(\lambda - 1)^2(\lambda - 8) = 0.$$

El valor $\lambda = 1$ da lugar a la solución $u + v + w = 0$, es decir, el punto E ha de estar en la recta del infinito. Para $\lambda = 8$ se obtiene que $E(1 : 2 : 4)$ y, por consiguiente, $F = A + E$ es $(7 : 0 : 0) + (1 : 2 : 4) = (4 : 1 : 2)$ y $D = B + F = (0 : 7 : 0) + (4 : 1 : 2) = (2 : 4 : 1)$.



TEMA 4
Aplicaciones afines

Como la estructura de espacio afín está directamente relacionada con la del espacio vectorial al cual está asociado, aplicaciones que conserven propiedades afines (como, por ejemplo, que conserven el baricentro) lo estarán con las aplicaciones lineales, que conservan las estructuras vectoriales.

4.1 Definición y propiedades de aplicaciones afines

4.1 Definición.- *Una aplicación $f\colon(\mathcal{A},\boldsymbol{E},\psi)\to(\mathcal{A}',\boldsymbol{E}',\psi')$ entre espacios afines se dice que es una* **aplicación afín** *si existe una aplicación lineal $\tilde f\colon\boldsymbol{E}\to\boldsymbol{E}'$ tal que, $\forall P,Q\in\mathcal{A}_1$:*

$$\tilde f(\overrightarrow{PQ})=\overrightarrow{f(P)f(Q)} \tag{4-1}$$

A la aplicación $\tilde f$ se le llama **aplicación lineal asociada** *a f.*

$$\begin{tikzcd}
\mathcal{A}\times\mathcal{A} \arrow[r, "f\times f"] \arrow[d, "\psi"'] & \mathcal{A}'\times\mathcal{A}' \arrow[d, "\psi'"] \\
\boldsymbol{E} \arrow[r, "\tilde f"] & \boldsymbol{E}'
\end{tikzcd}
\qquad
\begin{tikzcd}
(P,Q) \arrow[r, mapsto, "f\times f"] \arrow[d, "\psi"'] & \big(f(A),f(B)\big) \arrow[d, "\psi'"] \\
\overrightarrow{PQ} \arrow[r, mapsto, "\tilde f"] & \tilde f(\overrightarrow{PQ})=\overrightarrow{f(P)f(Q)}
\end{tikzcd}$$

Obsérvese que **una aplicación afín $f\colon\mathcal{A}\to\mathcal{A}'$ se determina conociendo la imagen de un punto $P\in\mathcal{A}$ y su aplicación lineal asociada $\tilde f$.**

En efecto, dado un punto $X\in\mathcal{A}$, su imagen $f(X)\in\mathcal{A}'$ queda determinada por P, $f(P)$ y $\tilde f$, pues como $\overrightarrow{PX}\in\boldsymbol{E}$ y por la definición de aplicación afín, $\tilde f(\overrightarrow{PX})=\overrightarrow{f(P)f(X)}$ y por tanto

$$f(X)=f(P)+\tilde f(\overrightarrow{PX}).$$

Podemos expresar la condición para que una aplicación entre espacios afines sea una aplicación afín de las siguientes formas:

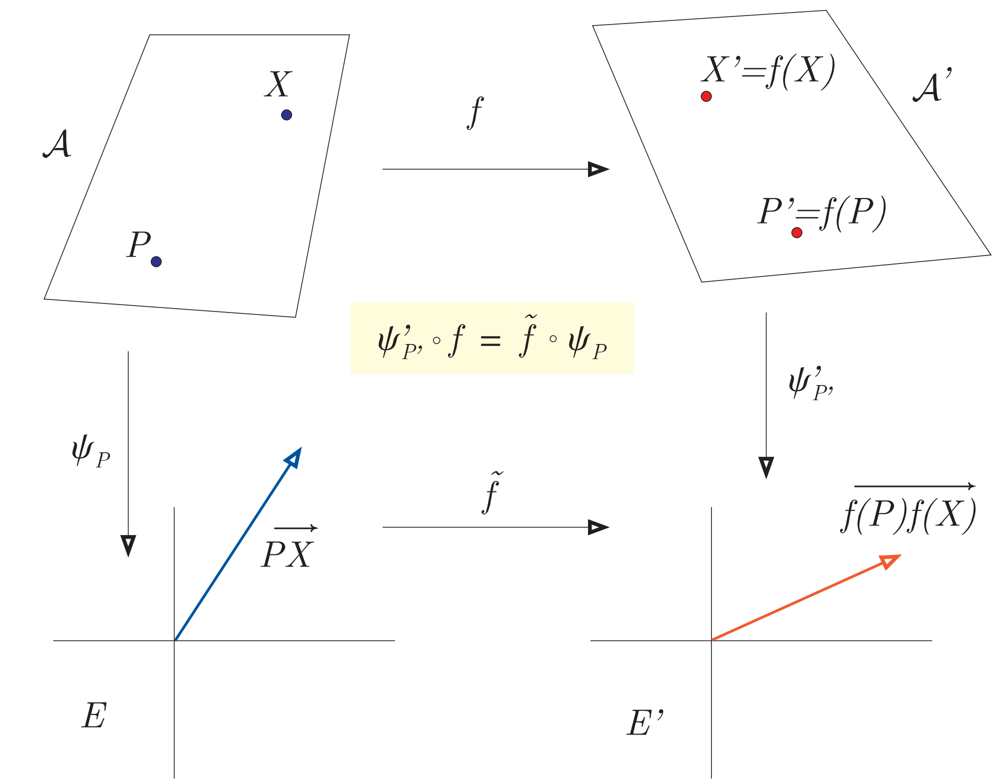
4.1' Definición.- *Una aplicación $f\colon \mathcal{A} \to \mathcal{A}$' es una aplicación afín si existe una aplicación lineal $\tilde{f}\colon \mathbf{E} \to \mathbf{E}$' tal que, $\forall P \in \mathcal{A}$ y $\forall \vec{\boldsymbol{v}} \in \mathbf{E}$:*

$$f(P + \vec{\boldsymbol{v}}) = f(P) + \tilde{f}(\vec{\boldsymbol{v}}). \tag{4-2}$$

Para simplificar la notación usamos el mismo símbolo "+" en ambos espacios afines.

4.1" Definición.- *Una aplicación $f\colon \mathcal{A} \to \mathcal{A}$' es una aplicación afín si existe una aplicación lineal $\tilde{f}\colon \mathbf{E} \to \mathbf{E}$' y existe un $P \in \mathcal{A}$ tal que el diagrama siguiente es conmutativo:*

$$\begin{tikzcd}
\mathcal{A} \arrow[r, mapsto, "f"] \arrow[d, "\psi_P"'] & \mathcal{A}' \arrow[d, "\psi'_{f(P)}"] \\
\mathbf{E} \arrow[r, mapsto, "\tilde{f}"] & \mathbf{E}'
\end{tikzcd}
\qquad
\begin{tikzcd}
X \arrow[r, mapsto, "f"] \arrow[d, "\psi_P"'] & f(X) \arrow[d, "\psi'_{f(P)}"] \\
\overrightarrow{PX} \arrow[r, mapsto, "\tilde{f}"] & \tilde{f}(\overrightarrow{PX}) = \overrightarrow{f(P)f(X)}
\end{tikzcd}$$



Un primer ejemplo de aplicaciones afines son las traslaciones ya introducidas en la tercera de las definiciones de espacio afín, en la Definición de la pág 19.

4.2 Ejemplo.- *En un espacio afín $(\mathcal{A}, \mathbf{E}, +)$ una* **traslación** *de vector $\vec{\boldsymbol{v}}$ es la aplicación $t_{\vec{\boldsymbol{u}}}\colon \mathcal{A} \to \mathcal{A}$ definida por $t_{\vec{\boldsymbol{v}}}(X) = X + \vec{\boldsymbol{v}}$, para todo $X \in \mathcal{A}$. Una traslación es una aplicación afín con aplicación lineal asociada la identidad $1_{\mathbf{E}}\colon \mathbf{E} \to \mathbf{E}$.*

Si $R = t_{\vec{\boldsymbol{v}}}(P) = P + \vec{\boldsymbol{v}}$ y $S = t_{\vec{\boldsymbol{v}}}(Q) = Q + \vec{\boldsymbol{v}}$, es decir, $\vec{\boldsymbol{v}} = \overrightarrow{PR} = \overrightarrow{QS}$, por la ley del paralelogramo (Proposición 2.3), se tiene que $\overrightarrow{PQ} = \overrightarrow{RS}$. Así, la aplicación lineal asociada $\tilde{t}_{\vec{\boldsymbol{v}}}$ es la identidad $1_{\mathbf{E}}$, pues:

$$\tilde{t}_{\vec{\boldsymbol{v}}}(\overrightarrow{PQ}) = \overrightarrow{t_{\vec{\boldsymbol{v}}}(P)t_{\vec{\boldsymbol{v}}}(Q)} = \overrightarrow{RS} = \overrightarrow{PQ}.$$

De hecho, se tiene que toda aplicación afín $f: \mathcal{A} \to \mathcal{A}$ cuya aplicación lineal asociada es la identidad $1_{\boldsymbol{E}}: \boldsymbol{E} \to \boldsymbol{E}$ es una traslación.

En efecto, tomemos un punto arbitrario $P \in \mathcal{A}$ y veamos que f es la traslación de vector $\vec{\boldsymbol{v}} = \overrightarrow{Pf(P)}$.

Como, para todo $X \in \mathcal{A}$, se tiene que $\overrightarrow{f(P)f(X)} = \tilde{f}(\overrightarrow{PX}) = 1_{\boldsymbol{E}}(\overrightarrow{PX}) = \overrightarrow{PX}$. De aquí, por la ley del paralelogramo, se sigue que $\overrightarrow{Xf(X)} = \overrightarrow{Pf(P)}$; o sea $f(X) = X + \overrightarrow{Pf(P)}$, entonces f es la traslación de vector $\vec{\boldsymbol{v}} = \overrightarrow{Pf(P)}$.

En los dos ejemplos siguientes veremos el comportamiento de las aplicaciones afines en el caso de considerar la estructura canónica de espacio afín sobre un espacio vectorial.

4.3 Ejemplo.- *Sean* $\boldsymbol{E}$ *y* $\boldsymbol{E}$'' *espacios vectoriales considerados como espacios afines* (Ejemplo 2.4), *una aplicación afín es la composición de una aplicación lineal con una traslación.*

Sean $f: \boldsymbol{E} \to \boldsymbol{E}$' una aplicación afín, $\tilde{f}: \boldsymbol{E} \to \boldsymbol{E}$' su aplicación lineal asociada y $\vec{\boldsymbol{w}} = f(\vec{o}_E)$. De la conmutatividad de los siguientes diagramas

$$\begin{array}{ccc} \boldsymbol{E} & \xmapsto{f} & \boldsymbol{E}\text{'} \\ \psi_{\vec{\boldsymbol{o}}} \downarrow & & \downarrow \psi\text{'}_{\vec{\boldsymbol{w}}} \\ \boldsymbol{E} & \xmapsto{\tilde{f}} & \boldsymbol{E}\text{'} \end{array} \qquad\qquad \begin{array}{ccc} \vec{\boldsymbol{u}} & \xmapsto{f} & f(\vec{\boldsymbol{u}}) \\ \psi_{\vec{\boldsymbol{o}}} \downarrow & & \downarrow \psi\text{'}_{\vec{\boldsymbol{w}}} \\ \vec{\boldsymbol{u}} & \xmapsto{\tilde{f}} & \tilde{f}(\vec{\boldsymbol{u}}) = f(\vec{\boldsymbol{u}}) - \vec{\boldsymbol{w}} \end{array}$$

surge que $f(\vec{\boldsymbol{u}}) = \tilde{f}(\vec{\boldsymbol{u}}) + \vec{\boldsymbol{w}}$; es decir:

$$f = t_{\vec{\boldsymbol{w}}} \circ \tilde{f}.$$

Obsérvese además que:

• Las traslaciones $t_{\vec{\boldsymbol{w}}}: \boldsymbol{E}\text{'} \to \boldsymbol{E}\text{'}$, $\vec{\boldsymbol{v}} \mapsto t_{\vec{\boldsymbol{w}}}(\vec{\boldsymbol{v}}) = \vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}$ no son aplicaciones lineales, a menos que $\vec{\boldsymbol{w}} = \vec{\boldsymbol{o}}$. Es decir, cuando la traslación sea la aplicación identidad en $\boldsymbol{E}$'.

• El modo de expresar una aplicación afín $f: \boldsymbol{E} \to \boldsymbol{E}$' como composición $f = t_{\vec{\boldsymbol{w}}} \circ \tilde{f}$, de una aplicación lineal con una traslación es único. Pues si $f = t_{\vec{\boldsymbol{z}}} \circ \hat{f}$, para $\vec{\boldsymbol{z}} \in \boldsymbol{E}$' y $\hat{f} \in \mathcal{L}_K(\boldsymbol{E}, \boldsymbol{E}\text{'})$, se tiene que $\vec{\boldsymbol{w}} = f(\vec{\boldsymbol{o}}) = t_{\vec{\boldsymbol{z}}}(\hat{f}(\vec{\boldsymbol{o}})) = t_{\vec{\boldsymbol{z}}}(\vec{\boldsymbol{o}}) = \vec{\boldsymbol{z}}$ y $f(\vec{\boldsymbol{v}}) = \tilde{f}(\vec{\boldsymbol{v}}) + \vec{\boldsymbol{w}} = \hat{f}(\vec{\boldsymbol{v}}) + \vec{\boldsymbol{w}}$, para todo $\vec{\boldsymbol{v}} \in \boldsymbol{E}$. Por tanto, $\hat{f} = \tilde{f}$.

• La composición, $h \circ t_{\vec{\boldsymbol{u}}}$, de una traslación $t_{\vec{\boldsymbol{u}}}: \boldsymbol{E} \to \boldsymbol{E}$ con cualquier aplicación lineal $h: \boldsymbol{E} \to \boldsymbol{E}$', es una aplicación afín; pues, para todo $\vec{\boldsymbol{v}} \in \boldsymbol{E}$ se tiene:

$$(h \circ t_{\vec{\boldsymbol{u}}})(\vec{\boldsymbol{v}}) = h(\vec{\boldsymbol{v}} + \vec{\boldsymbol{u}}) = h(\vec{\boldsymbol{v}}) + h(\vec{\boldsymbol{u}}) = (t_{h(\vec{\boldsymbol{u}})} \circ h)(\vec{\boldsymbol{v}}).$$

Luego, $h \circ t_{\vec{\boldsymbol{u}}}$ es una aplicación afín de aplicación lineal asociada h.

• Si $f: \boldsymbol{E} \to \boldsymbol{E}$' es una aplicación afín biyectiva existe una única aplicación lineal $h: \boldsymbol{E} \to \boldsymbol{E}$' y un único vector $\vec{\boldsymbol{u}} \in \boldsymbol{E}$ tal que $f = h \circ t_{\vec{\boldsymbol{u}}}$.

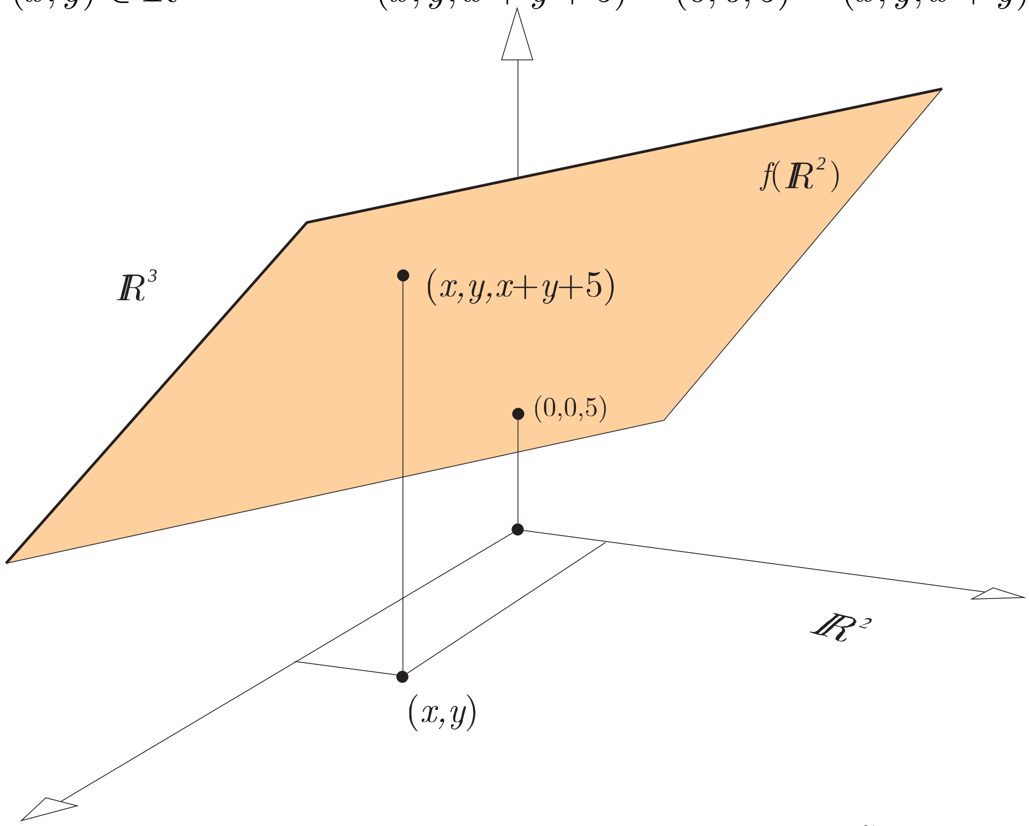
En efecto, como existe un único $\vec{\boldsymbol{w}} \in \boldsymbol{E}$' y una única aplicación $h\colon \boldsymbol{E} \to \boldsymbol{E}$', tal que $f = t_{\vec{\boldsymbol{w}}} \circ h$ y como f y $t_{\vec{\boldsymbol{w}}}$ son biyectivas, también h es biyectiva y, por tanto, existe un único $\vec{\boldsymbol{u}} \in \boldsymbol{E}$ tal que $\vec{\boldsymbol{w}} = h(\vec{\boldsymbol{u}})$, y se tiene que

$$f = t_{\vec{\boldsymbol{w}}} \circ h = t_{h(\vec{\boldsymbol{u}})} \circ h = h \circ t_{\vec{\boldsymbol{u}}}.$$

Siguiendo con el caso particular de espacios vectoriales considerados con espacios afines, consideremos la siguiente situación:

4.3' Ejemplo.- *Sea la aplicación $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ entre los espacios afines $\mathbb{R}^2$ y $\mathbb{R}^3$, definida por $(x,y) \mapsto (x,y,x+y+5)$.*

$$\begin{tikzcd}
(x,y) \in \mathbb{R}^2 \arrow[r, mapsto, "f"] \arrow[d, "\psi_{(0,0)}"'] & (x,y,x+y+5) \in \mathbb{R}^3 \arrow[d, "\psi_{(0,0,5)}"] \\
(x,y) \in \mathbb{R}^2 \arrow[r, mapsto, "\tilde{f}"] & (x,y,x+y+5) - (0,0,5) = (x,y,x+y) \in \mathbb{R}^3
\end{tikzcd}$$



La aplicación afín f es composición de la aplicación lineal $\tilde{f}$ con la traslación $t_{(0,0,5)}(x,y,z) = (x,y,z+5)$.

En este ejemplo se observa que la imagen de la aplicación afín f, $f(\mathbb{R}^2)$, es un plano (variedad lineal) de $\mathbb{R}^3$. Este hecho es una situación particular de un resultado más general, a saber:

4.4 Proposición.- *Las aplicaciones afines conservan los subespacios afines, es decir, si $f\colon \mathcal{A} \to \mathcal{A}$' es una aplicación afín con aplicación lineal asociada $\tilde{f}\colon \boldsymbol{E} \to \boldsymbol{E}$' y $\mathfrak{F} = P + \boldsymbol{F}$ es una variedad lineal afín de $\boldsymbol{E}$ de subespacio director $\boldsymbol{F}$, entonces $f(\mathfrak{F}) = f(P) + \tilde{f}(\boldsymbol{F})$ es una variedad lineal afín de $\mathcal{A}$' de subespacio director $\tilde{f}(\boldsymbol{F})$.*

Demostración.- $\tilde{f}(\boldsymbol{F})$ es un subespacio vectorial de $\boldsymbol{E}$' (Proposición 1.34).

Para todo $\vec{\boldsymbol{u}} \in \boldsymbol{F}$, $P + \vec{\boldsymbol{u}} \in \mathcal{F}$ y $f(P + \vec{\boldsymbol{u}}) = f(P) + \tilde{f}(\vec{\boldsymbol{u}}) \in f(P) + \tilde{f}(\boldsymbol{F})$.

Sea ahora $Q \in f(P) + \tilde{f}(\boldsymbol{F})$, entonces $\overrightarrow{f(P)Q} \in \tilde{f}(\boldsymbol{F})$ y $\exists \vec{\boldsymbol{u}} \in \boldsymbol{F}$ tal que $\tilde{f}(\vec{\boldsymbol{u}}) = \overrightarrow{f(P)Q}$; luego, $f(P + \vec{\boldsymbol{u}}) = f(P) + \tilde{f}(\vec{\boldsymbol{u}}) = Q$ y, por tanto, $Q \in f(\mathcal{F})$. ◼

4.5 Corolario.- *Las imágenes de dos variedades lineales paralelas mediante una aplicación afín son variedades lineales paralelas.*

Demostración.- Sea $f\colon \mathcal{A} \to \mathcal{A}$' una aplicación afín entre espacios afines con aplicación lineal asociada $\tilde{f}\colon \boldsymbol{E} \to \boldsymbol{E}$', si $P + \boldsymbol{F}$ y $Q + \boldsymbol{G}$ son variedades lineales paralelas, el subespacio director de una de ellas está contenido en el de la otra. Sea, por ejemplo, $\boldsymbol{F} \subseteq \boldsymbol{G}$, entonces $\tilde{f}(\boldsymbol{F}) \subseteq \tilde{f}(\boldsymbol{G})$. Por consiguiente, $f(P) + \tilde{f}(\boldsymbol{F})$ y $f(Q) + \tilde{f}(\boldsymbol{G})$ son variedades lineales de $\mathcal{A}$' paralelas. ◼

4.6 Nota.- Obsérvese que de esta proposición y su corolario, se sigue que la imagen, mediante una aplicación afín, de una recta es un punto o una recta; y que la imagen de un plano es o bien un punto, una recta o un plano. En consecuencia, la imagen de puntos alineados, mediante una aplicación afín, son puntos alineados (pudiendo ocurrir que estos puntos se reduzcan a uno solo, si la aplicación no es inyectiva).
También se tiene que la imagen de dos rectas paralelas son rectas (si la aplicación afín es inyectiva) paralelas.

4.7 Proposición.- *Sean $(\mathcal{A}, \boldsymbol{E})$ un espacio afín de dimensión finita, una aplicación afín $f\colon \mathcal{A} \to \mathcal{A}$ y $\mathcal{F} = \{X \in \mathcal{A} / f(X) = X\}$ el conjunto de puntos fijos de f.*

— Si $\mathcal{F} \neq \varnothing$, entonces $\mathcal{F}$ es una variedad lineal de subespacio director $\mathrm{Ker}(\tilde{f} - 1_{\boldsymbol{E}})$, esto es, el subespacio vectorial de vectores propio correspondiente al valor propio $1_{I\!R}$ de la aplicación lineal asociada a f.

— f tienen único punto fijo $\Leftrightarrow \mathrm{Ker}(\tilde{f} - 1_{\boldsymbol{E}}) = \{\vec{\boldsymbol{0}}_{\boldsymbol{E}}\}$.

Demostración.- Si P es un punto fijo de f, para todo $X \in \mathcal{A}$ se tiene que $\tilde{f}(\overrightarrow{PX}) = \overrightarrow{f(P)f(X)} = \overrightarrow{Pf(X)}$. Así,

$$X \in \mathcal{F} \Leftrightarrow f(X) = X \Leftrightarrow \tilde{f}(\overrightarrow{PX}) = \overrightarrow{PX} \Leftrightarrow \overrightarrow{PX} \in \mathrm{Ker}(\tilde{f} - 1_{\boldsymbol{E}}).$$

Luego, el conjunto de puntos fijos de f es una variedad lineal de subespacio director $\mathrm{Ker}(\tilde{f} - 1_{\boldsymbol{E}})$:

$$\mathcal{F} = \{X \in \mathcal{A} / f(X) = X\} = P + \mathrm{Ker}(\tilde{f} - 1_{\boldsymbol{E}}).$$

Para demostrar la segunda parte, supongamos que f tiene un único punto fijo P, entonces

$$(\tilde{f} - 1_{\boldsymbol{E}})(\overrightarrow{PX}) = \vec{\boldsymbol{0}}_{\boldsymbol{E}} \Rightarrow \overrightarrow{f(P)f(X)} - \overrightarrow{PX} = \overrightarrow{Pf(X)} - \overrightarrow{PX} = \vec{\boldsymbol{0}}_{\boldsymbol{E}} \Rightarrow f(X) = X.$$

Al ser P es el único punto fijo, $X = P$. Es decir, $\mathrm{Ker}(\tilde{f} - 1_{\boldsymbol{E}}) = \{\vec{\boldsymbol{0}}_{\boldsymbol{E}}\}$.

Supongamos ahora que $\mathrm{Ker}(\tilde{f} - 1_{\boldsymbol{E}}) = \{\vec{\boldsymbol{0}}_{\boldsymbol{E}}\}$, es decir, que $\tilde{f} - 1_{\boldsymbol{E}}$ es inyectiva. Como E es de dimensión finita $\tilde{f} - 1_{\boldsymbol{E}}$ es suprayectiva.

Sea $P_0 \in \mathcal{A}$, para todo $X \in \mathcal{A}$ podemos poner
$\overrightarrow{Xf(X)} = \overrightarrow{P_0f(P_0)} + \overrightarrow{f(P_0)f(X)} - \overrightarrow{P_0X} = \overrightarrow{P_0f(P_0)} + \tilde{f}(\overrightarrow{P_0X}) - \overrightarrow{P_0X} =$

$$= \overrightarrow{P_0f(P_0)} + (\tilde{f} - 1_{\boldsymbol{E}})(\overrightarrow{P_0X}).$$

Así, un punto fijo X de f ha de satisfacer a:

$$(\tilde{f} - 1_{\boldsymbol{E}})(\overrightarrow{P_0X}) = -\overrightarrow{P_0f(P_0)}.$$

Como $\tilde{f} - 1_{\boldsymbol{E}}$ es suprayectiva, existe al menos un $P \in \mathcal{A}$, que satisface a esta relación; y como es inyectiva P es único. ◼

Damos a continuación dos caracterizaciones de aplicaciones afines en términos de baricentro y razón simple que pueden ser tomadas como definiciones de aplicación afines en un sentido más intrínseco, ya que sólo involucran puntos de los espacios afines.

4.8 Proposición[Definición].- *Consideremos dos espacios afines $(\mathcal{A}, \boldsymbol{E}, +)$ y $(\mathcal{A}', \boldsymbol{E}', +)$ y una aplicación $f\colon \mathcal{A} \to \mathcal{A}'$, entonces:*
f es una aplicación afín si y sólo si para toda familia de puntos ponderados $\{(P_i, \alpha^i)\}_{i\in I}$ de $\mathcal{A}$ tal que $\sum_{i\in I} \alpha^i = 1_K$, se tiene que

$$f\Big(\sum_{i\in I} \alpha^i P_i\Big) = \sum_{i\in I} \alpha^i f(P_i).$$

(ver notación (3-6)). Es decir, $f\colon \mathcal{A} \to \mathcal{A}'$ es una aplicación afín si y sólo si conserva el baricentro.

Demostración.- Sea $\tilde{f}\colon \boldsymbol{E} \to \boldsymbol{E}'$ la aplicación lineal asociada a la aplicación afín f.

Si B es el baricentro de la familia de puntos ponderados de $\mathcal{A}$, $\{(P_i, \alpha^i)\}_{i\in I}$, con $\sum_{i\in I} \alpha^i = 1_K$, se tiene que (Definición de la página 40):

$$\sum_{i\in I} \alpha^i \overrightarrow{BP_i} = \vec{\boldsymbol{0}}_{\boldsymbol{E}}.$$

Luego,

$$\vec{\boldsymbol{0}}_{\boldsymbol{E}'} = \tilde{f}\Big(\sum_{i\in I} \alpha^i \overrightarrow{BP_i}\Big) = \sum_{i\in I} \alpha^i \tilde{f}(\overrightarrow{BP_i}) = \sum_{i\in I} \alpha^i \overrightarrow{f(B)f(P_i)}.$$

Con los que $f(B)$ es el baricentro de los puntos $\{(f(P_i), \alpha^i)\}_{i\in I}$ de $\mathcal{A}'$.

Para demostrar la otra implicación, tomemos un punto $P \in \mathcal{A}$ y definimos la aplicación

$$\hat{f}\colon \boldsymbol{E} \to \boldsymbol{E}' \qquad \vec{\boldsymbol{v}} \mapsto \hat{f}(\vec{\boldsymbol{v}}) = \overrightarrow{f(P)f(P + \vec{\boldsymbol{v}})}.$$

Demostremos que $\hat{f}$ es la aplicación lineal asociada a f.

Para demostrar que $\hat{f}$ es lineal, sean $\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \in \boldsymbol{E}$ y $\lambda, \mu \in K$.

$$P + \lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}} = P + \lambda\overrightarrow{P(P + \vec{\boldsymbol{u}})} + \mu\overrightarrow{P(P + \vec{\boldsymbol{v}})} + (1 - (\lambda + \mu))\overrightarrow{PP}.$$

Luego, $P + \lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}$ es el baricentro de los puntos ponderados $(P + \vec{\boldsymbol{u}}, \lambda)$, $(P + \vec{\boldsymbol{v}}, \mu)$, $(P, 1 - (\lambda + \mu))$.

Como f conserva el baricentro, $f(P+\lambda\vec{\boldsymbol{u}}+\mu\vec{\boldsymbol{v}})$ es el baricentro de los puntos $(f(P + \vec{\boldsymbol{u}}), \lambda), (f(P + \vec{\boldsymbol{v}}), \mu), (f(P), 1 - (\lambda + \mu))$.

Se tiene, por tanto, que

$$\overrightarrow{f(P)f(P+\lambda\vec{\boldsymbol{u}}+\mu\vec{\boldsymbol{v}})} = \lambda\overrightarrow{f(P)f(P+\vec{\boldsymbol{u}})} + \mu\overrightarrow{f(P)f(P+\vec{\boldsymbol{v}})} + (1 - \lambda - \mu)\overrightarrow{f(P)f(P)}$$

Es decir,
$$\hat{f}(\lambda\vec{\boldsymbol{u}}+\mu\vec{\boldsymbol{v}})=\lambda\hat{f}(\vec{\boldsymbol{u}})+\mu\hat{f}(\vec{\boldsymbol{v}}).$$

Se concluye que $\hat{f}$ es lineal y por su propia construcción, resulta que f es una aplicación afín con aplicación lineal asociada $\tilde{f}=\hat{f}$ ◼

4.9 Nota.- La aplicación $\hat{f}$ definida en la demostración de la proposición anterior no depende del punto P tomado.

En efecto, sea otro punto $Q\in\mathcal{A}$, entonces
$$Q+\vec{\boldsymbol{v}}=P+\overrightarrow{PQ}+\vec{\boldsymbol{v}}=P-\overrightarrow{PP}+\overrightarrow{PQ}+\overrightarrow{P(P+\vec{\boldsymbol{v}})}.$$
Así, $Q+\vec{\boldsymbol{v}}$ es baricentro de los puntos ponderados $(P,-1),(Q,1),(P+\vec{\boldsymbol{v}},1)$.

Como f conserva el baricentro, resulta que $f(Q+\vec{\boldsymbol{v}})$ es el baricentro de los puntos $(f(P),-1),(f(Q),1),(f(P+\vec{\boldsymbol{v}}),1)$; entonces
$$\overrightarrow{f(Q)f(Q+\vec{\boldsymbol{v}})}=-\overrightarrow{f(Q)f(P)}+\overrightarrow{f(Q)f(Q)}+\overrightarrow{f(Q)f(P+\vec{\boldsymbol{v}})}=\overrightarrow{f(P)f(P+\vec{\boldsymbol{v}})}.$$

Para dar una caracterización de aplicaciones afines en términos de conservación de la razón simple, necesitamos que la aplicación sea al menos inyectiva, para que la imagen de puntos distintos sean también diferentes.

4.10 Proposición.- *Sea $f\colon\mathcal{A}\to\mathcal{A}$' una aplicación biyectiva entre espacios afines, asociados a espacios vectoriales $\boldsymbol{E}$ y $\boldsymbol{E}$', sobre un mismo cuerpo K de característica distinta de 2. Entonces:*
f es una aplicación afín $\Leftrightarrow$ f aplica puntos alineados en puntos alineados y conserva la razón simple.

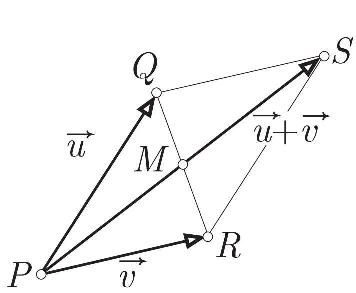
Demostración.- ($\Rightarrow$) Por la Nota 4.6, si f es una aplicación afín biyectiva lleva puntos alineados en puntos alineados. Además, si P_1,P_2 y P_3 ($P_1\neq P_2$) son tres puntos alineados, la razón simple (Definición 3.25) $(P_1\ P_2\ P_3)$ es el escalar tal que $\overrightarrow{P_1P_3}=(P_1\ P_2\ P_3)\ \overrightarrow{P_1P_2}$.

Esto quiere decir que P_1 es el baricentro de P_2 y P_3 afectados de los "pesos" $-(P_1\ P_2\ P_3)$ y 1, respectivamente. Y como f conserva el baricentro, $f(P_1)$ es el baricentro de $f(P_2)$ y $f(P_3)$ afectados de los "pesos" $-(P_1\ P_2\ P_3)$ y 1, respectivamente; esto es,
$$\overrightarrow{f(P_1)f(P_3)}-(P_1\ P_2\ P_3)\ \overrightarrow{f(P_1)f(P_2)}=\vec{\boldsymbol{0}}\Leftrightarrow\overrightarrow{f(P_1)f(P_3)}=(P_1\ P_2\ P_3)\ \overrightarrow{f(P_1)f(P_2)},$$
por lo que $(f(P_1)\ f(P_2)\ f(P_3))=(P_1\ P_2\ P_3)$.

($\Leftarrow$) Sea $P\in\mathcal{A}$ y $\hat{f}\colon\boldsymbol{E}\to\boldsymbol{E}$' definida, para $\vec{\boldsymbol{v}}\in\boldsymbol{E}$ por $\hat{f}(\vec{\boldsymbol{v}})=\overrightarrow{f(P)f(P+\vec{\boldsymbol{v}})}$.

Si demostramos que $\hat{f}$ es lineal, entonces f es una aplicación afín, determinada por $\hat{f}$, el punto P y su imagen $f(P)\in\mathcal{A}$'. La aplicación lineal asociada es, entonces, $\tilde{f}=\hat{f}$.

Para $\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}\in\boldsymbol{E}$, existen $Q,R,S\in\mathcal{A}$ tales que $\vec{\boldsymbol{u}}=\overrightarrow{PQ}$, $\vec{\boldsymbol{v}}=\overrightarrow{PR}$ y $\vec{\boldsymbol{u}}+\vec{\boldsymbol{v}}=\overrightarrow{PS}$. Luego, $\overrightarrow{PS}=\overrightarrow{PQ}+\overrightarrow{PR}=\overrightarrow{PQ}+\overrightarrow{QS}$, entonces $\overrightarrow{PR}=\overrightarrow{QS}$.



Si M es el punto medio de P y S, se tiene que $\overrightarrow{PM}= 1/2\overrightarrow{PS}$ (el cuerpo K es de característica distinta de 2). Si N es el punto medio de Q y R, N es el baricentro de Q y R afectados del mismo "pesos", es decir,

$$\overrightarrow{PN} = \frac{1}{2}\overrightarrow{PQ} + \frac{1}{2}\overrightarrow{PR} = \frac{1}{2}\overrightarrow{PQ} + \frac{1}{2}\overrightarrow{QS} = \frac{1}{2}\overrightarrow{PS}.$$

Se concluye que $\overrightarrow{PM}= \overrightarrow{PN}$, luego $M = N$.

Por otra parte, en términos de la razón simple, que M y N sean los puntos medios de P y S y de Q y R, respectivamente, significa que

$$(P\,S\,M) = \frac{1}{2} \qquad \text{y} \qquad (Q\,R\,N) = \frac{1}{2}.$$

Así, por la hipótesis sobre f, $f(P), f(S)$ y $f(M)$ son distintos y están alineados y $(f(P)\,f(S)\,f(M)) = \frac{1}{2}$. Análogamente, $(f(Q)\,f(R)\,f(N)) = \frac{1}{2}$.

Como $M = N$, se tiene $f(M) = f(N)$; con lo que $f(M)$ es el punto medio de $f(P)$ y $f(S)$, y también es el punto medio de $f(Q)$ y $f(R)$; con lo que $\overrightarrow{f(P)f(S)} = 2\overrightarrow{f(P)f(M)}$ y $\overrightarrow{f(Q)f(R)} = 2\overrightarrow{f(M)f(R)}$ Luego:

$$\overrightarrow{f(P)f(S)} = 2\overrightarrow{f(P)f(M)} = 2\overrightarrow{f(P)f(Q)} + 2\overrightarrow{f(Q)f(M)},$$
$$\overrightarrow{f(P)f(S)} = 2\overrightarrow{f(P)f(M)} = 2\overrightarrow{f(P)f(R)} + 2\overrightarrow{f(R)f(M)}.$$

Sumando miembro a miembro, teniendo en cuenta que $f(M)$ es el punto medio de $f(Q)$ y $f(R)$ y simplificando, se obtiene

$$\overrightarrow{f(P)f(S)} = \overrightarrow{f(P)f(Q)} + \overrightarrow{f(P)f(R)},$$

Es decir,

$$\overrightarrow{f(P)f(P+\vec{\boldsymbol{u}}+\vec{\boldsymbol{v}})} = \overrightarrow{f(P)f(P+\vec{\boldsymbol{u}})} + \overrightarrow{f(P)f(P+\vec{\boldsymbol{v}})}.$$
$$\hat{f}(\vec{\boldsymbol{u}}+\vec{\boldsymbol{v}}) = \hat{f}(\vec{\boldsymbol{u}}) + \hat{f}(\vec{\boldsymbol{v}}).$$

Sean ahora $\vec{\boldsymbol{u}} \in \mathbf{E}$ y $\lambda \in K$, existen $Q, R \in \mathcal{A}$ tal que $\vec{\boldsymbol{u}} = \overrightarrow{PQ}, \lambda\vec{\boldsymbol{u}} = \overrightarrow{PR}$, entonces,

$$(P\,Q\,R) = \lambda \quad \text{y} \quad (f(P)\,f(Q)\,f(R)) = \lambda.$$
$$\overrightarrow{f(P)f(R)} = \lambda\overrightarrow{f(P)f(Q)} \Rightarrow \overrightarrow{f(P)f(P+\lambda\vec{\boldsymbol{u}})} = \lambda\overrightarrow{f(P)f(Q+\vec{\boldsymbol{u}})}.$$
$$\hat{h}(\lambda\vec{\boldsymbol{u}}) = \lambda\hat{f}(\vec{\boldsymbol{u}}).$$

Se concluye que $\hat{f}$ es lineal. ■

La demostración de la siguiente caracterización de aplicaciones afines, sólo requiriendo que sea biyectiva y que transformen rectas en rectas, puede consultarse en [9, pag. 76]; para el caso más general de espacios proyectivos en [6, pag. 104] o para planos proyectivos en [5, Apéndice B].

4.11 Proposición[Teorema fundamental de la geometría afín].- *Dada una aplicación biyectiva $f\colon \mathcal{A} \to \mathcal{A}'$ entre espacios afines de la misma dimensión $n \geq 0$, entonces:*
f es una aplicación afín si y sólo si f aplica puntos alineados en puntos alineados. ■

4.2 Grupos de afinidades

La composición de aplicaciones afines es una aplicación afín:

4.12 Proposición.- *Sean $(\mathcal{A}_i, E_i)$, $i = 1, 2, 3$, espacios afines, $f\colon \mathcal{A}_1 \to \mathcal{A}_2$ y $g\colon \mathcal{A}_2 \to \mathcal{A}_3$ aplicaciones afines de aplicaciones lineales asociadas $\tilde{f}\colon E_1 \to E_2$ y $\tilde{g}\colon E_2 \to E_3$. Entonces $g \circ f$ es una aplicación afín de aplicación lineal asociada $\widetilde{g \circ f} = \tilde{g} \circ \tilde{f}$.*

Demostración.- Basta con verificar que $\widetilde{g \circ f} = \tilde{g} \circ \tilde{f}$. Para todo $P, Q \in \mathcal{A}_1$ se tiene:

$$\widetilde{g \circ f}(\overrightarrow{PQ}) = \overrightarrow{(g \circ f)(P)(g \circ f)(Q)} = \overrightarrow{g(f(P))g(f(Q))} = \tilde{g}(\overrightarrow{f(P)f(Q)}) =$$

$$= \tilde{g}\big(\tilde{f}(\overrightarrow{PQ})\big) = (\tilde{g} \circ \tilde{f})(\overrightarrow{PQ}).$$

■

En particular, cuando consideramos los espacios vectoriales como espacios afines, veamos como son la aplicación lineal y la traslación que determinan la composición de dos aplicaciones afines (ver Ejemplo 4.3).

4.13 Proposición.- *Sean E_i $(i = 1, 2, 3)$ tres espacios vectoriales considerados como espacios afines, $f\colon E_1 \to E_2$ y $g\colon E_2 \to E_3$ dos aplicaciones afines, con $f = t_{f(\vec{0}_{E_1})} \circ \tilde{f}$ y $g = t_{g(\vec{0}_{E_2})} \circ \tilde{g}$. Entonces,*

$$g \circ f = t_{(g \circ f)(\vec{0})} \circ \widetilde{g \circ f}.$$

Demostración.-

$$g \circ f = (t_{g(\vec{0}_{E_2})} \circ \tilde{g}) \circ (t_{f(\vec{0}_{E_1})} \circ \tilde{f}) = t_{g(\vec{0}_{E_2})} \circ t_{\tilde{g}(f(\vec{0}_{E_1}))} \circ \tilde{g} \circ \tilde{f} =$$

$$= t_{g(\vec{0}_{E_2}) + \tilde{g}(f(\vec{0}_{E_1}))} \circ \widetilde{g \circ f} = t_{(g \circ f)(\vec{0}_{E_1})} \circ \widetilde{g \circ f}.$$

■

El vector de traslación es $(g \circ f)(\vec{0}_{E_1}) = g(\vec{0}_{E_2}) + \tilde{g}(f(\vec{0}_{E_1}))$.
Por lo que, la aplicación afín $g \circ f$ es lineal cuando $g(\vec{0}_{E_2}) = -\tilde{g}(f(\vec{0}_{E_1}))$.

La relación existente entre una aplicación afín y su aplicación lineal asociada se pone de manifiesto en el siguiente resultado:

4.14 Proposición.- *Dada una aplicación afín $f\colon \mathcal{A} \to \mathcal{A}'$ de aplicación lineal asociada $\tilde{f}\colon E \to E'$ se tiene que:*

i) f es inyectiva $\Leftrightarrow$ $\tilde{f}$ es inyectiva

ii) f es suprayectiva $\Leftrightarrow$ $\tilde{f}$ es suprayectiva

Demostración.- *i)* Como $\tilde{f}(\overrightarrow{PQ}) = \vec{0}_{E'} \Leftrightarrow \overrightarrow{f(P)f(Q)} = \vec{0}_{E'} \Leftrightarrow f(P) = f(Q)$.

$$f \text{ inyectiva y } \tilde{f}(\overrightarrow{PQ}) = \vec{0}_{E'} \Rightarrow P = Q \Rightarrow \overrightarrow{PQ} = \vec{0}_{E}.$$

$$\tilde{f} \text{ inyectiva y } f(P) = f(Q) \Rightarrow \overrightarrow{PQ} = \vec{\mathbf{0}}_E \Rightarrow P = Q.$$

ii) Supongamos que f es suprayectiva y $\vec{\boldsymbol{w}} \in \boldsymbol{E}$'. Existen dos puntos R y S en $\mathcal{A}$', tales que $\vec{\boldsymbol{w}} = \overrightarrow{RS}$ y dos puntos P y Q en $\mathcal{A}$ tales que $R = f(P)$ y $S = f(Q)$. Por tanto, $\vec{\boldsymbol{w}} = \overrightarrow{RS} = \overrightarrow{f(P)f(Q)} = \tilde{f}(\overrightarrow{PQ})$. Es decir $\tilde{f}$ es suprayectiva.

Si $\tilde{f}$ es suprayectiva y $R \in \mathcal{A}$', sean los puntos auxiliares $Q \in \mathcal{A}$ y $S = f(Q) \in f(\mathcal{A})$; existen $\vec{\boldsymbol{v}} \in \boldsymbol{E}$ y $P \in \mathcal{A}$ tales que $\vec{\boldsymbol{v}} = \overrightarrow{PQ}$ y $\tilde{f}(\vec{\boldsymbol{v}}) = \overrightarrow{RS}$. Así, $\overrightarrow{RS} = \tilde{f}(\vec{\boldsymbol{v}}) = \tilde{f}(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)} = \overrightarrow{f(P)S}$. Luego, $f(P) = R$ y f es sobreyectiva.

◪

4.15 Corolario.- *Dada una aplicación afín biyectiva $f: \mathcal{A} \to \mathcal{A}$' de aplicación lineal asociada (biyectiva) $\tilde{f}: \boldsymbol{E} \to \boldsymbol{E}$', se tiene que f^{-1} es una aplicación afín con aplicación lineal asociada $\widetilde{f^{-1}} = \tilde{f}^{-1}$.*

Demostración.-

$$\left.\begin{array}{l} 1_E = \widetilde{1_{\mathcal{A}}} = \widetilde{f^{-1} \circ f} = \widetilde{f^{-1}} \circ \tilde{f} \\ 1_{E'} = \widetilde{1_{\mathcal{A}'}} = \widetilde{f \circ f^{-1}} = \tilde{f} \circ \widetilde{f^{-1}} \end{array}\right\} \Rightarrow (\tilde{f})^{-1} = \widetilde{f^{-1}}.$$

4.16 Definición.- *Si una aplicación afín es biyectiva diremos que es un* ***isomorfismo afín*** *o* ***transformación afín****. Cuando el isomorfismo afín es sobre el mismo espacio afín se dice que es una* ***afinidad****. Dos espacio afines son* ***isomorfos*** *si entre ellos existe un isomorfismo afín.*

El conjunto de todas las afinidades de un espacio afín $\mathcal{A}$ es un grupo con la operación composición, que se suele denotar por $A(\mathcal{A})$.

4.17 Ejemplo.- *Existe una aplicación afín biyectiva entre todo espacio afín $\mathcal{A}$ asociado a un espacio vectorial $\boldsymbol{E}$ de dimensión finita n sobre un cuerpo K y el espacio afín canónico K^n* (Ejemplo 2.5). *Dicho de otra forma, todos los espacio afines de dimensión n sobre un cuerpo K son isomorfos al espacio afín canónico K^n.*

Tal aplicación afín biyectiva se obtiene dando una referencia cartesiana $\mathcal{R} = \{O; \vec{\boldsymbol{u}}_1, \ldots, \vec{\boldsymbol{u}}_n\}$ y definiendo

$$f: \mathcal{A} \to K^n, \quad P \mapsto (x^1, \ldots, x^n)$$

siendo $(x^1, \ldots, x^n)$ las coordenadas de P en $\mathcal{R}$.

Si Q es un punto de $\mathcal{A}$ de coordenadas $(y^1, \ldots, y^n)$, la aplicación lineal asociada a f está dada por:

$$\tilde{f}(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)} = (y^1 - x^1, \ldots, y^n - x^n).$$

$\tilde{f}$ es la aplicación lineal determinada por $\tilde{f}(\vec{\boldsymbol{u}}_i) = \vec{\boldsymbol{e}}_i$ $(i = 1, \ldots, n)$, siendo $\vec{\boldsymbol{e}}_i$ el i–ésimo vector de la base canónica de K^n.

En este caso, el grupo afín se denota por $GA(n, K)$, del cual el grupo general lineal $GL(n, K)$ es un subgrupo.

4.3 Expresión matricial de una aplicación afín

Usando referencias las aplicaciones afines se pueden representar en términos de matrices.

4.18 Proposición.- *Sean una aplicación afín $f\colon \mathcal{A} \to \mathcal{A}'$ de aplicación lineal asociada $\tilde{f}\colon \boldsymbol{E} \to \boldsymbol{E}'$, $\mathcal{R} = \{O; \vec{\boldsymbol{u}}_1, \ldots, \vec{\boldsymbol{u}}_n\}$ y $\mathcal{R}' = \{O'; \vec{\boldsymbol{v}}_1, \ldots, \vec{\boldsymbol{v}}_m\}$ referencias cartesianas en $\boldsymbol{E}$ y $\boldsymbol{E}'$, respectivamente. Si las coordenadas de un punto $X \in \mathcal{A}$ son $(x^1, \ldots, x^n)$ respecto de $\mathcal{R}$, las coordenadas de su imagen $f(X)$ son $(y^1, \ldots, y^m)$ respecto de $\mathcal{R}'$ y las coordenadas de $f(O)$ son $(b^1, \ldots, b^m)$ respecto de $\mathcal{R}'$, entonces*

$$\begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^m \end{pmatrix} + \begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m & a_2^m & \cdots & a_n^m \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}.$$

La matriz $A = (a_i^j)$ es la matriz asociada a la aplicación $\tilde{f}$ respecto de las bases $\{\vec{\boldsymbol{u}}_1, \ldots, \vec{\boldsymbol{u}}_n\}$ de $\boldsymbol{E}$ y $\{\vec{\boldsymbol{v}}_1, \ldots, \vec{\boldsymbol{v}}_m\}$ de $\boldsymbol{E}'$.

Demostración.- Usar la expresión matricial (dada en la Proposición 1.41) de una aplicación lineal, que $f(X) = f(O) + \tilde{f}(\overrightarrow{OX})$ y, por tanto, que

$$\overrightarrow{O'f(X)} = \overrightarrow{O'f(O)} + \tilde{f}(\overrightarrow{OX}).$$

Podemos poner las ecuaciones de una aplicación afín por

$$\begin{pmatrix} 1 \\ Y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ B & A \end{pmatrix} \begin{pmatrix} 1 \\ X \end{pmatrix},$$

siendo X e Y matrices columna formadas por las coordenadas cartesianas $(x^1, \ldots, x^n)$ y $(y^1, \ldots, y^m)$ de los puntos X y $f(X)$ respecto a las referencias $\mathcal{R}$ y $\mathcal{R}'$, respectivamente; B la matriz columna formada por las coordenadas cartesianas $(b^1, \ldots, b^m)$ del punto $f(O)$ respecto a la referencia $\mathcal{R}'$; $A = (a_i^j)$ es la matriz asociada a la aplicación $\tilde{f}$ respecto de las bases $\{\vec{\boldsymbol{u}}_1, \ldots, \vec{\boldsymbol{u}}_n\}$ de $\boldsymbol{E}$ y $\{\vec{\boldsymbol{v}}_1, \ldots, \vec{\boldsymbol{v}}_m\}$ de $\boldsymbol{E}'$; y donde 0 indica un matriz fila con n ceros:

$$\begin{pmatrix} 1 \\ y^1 \\ y^2 \\ \vdots \\ y^n \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ b^1 & a_1^1 & a_2^1 & \cdots & a_n^1 \\ b^2 & a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b^m & a_1^m & a_2^m & \cdots & a_n^m \end{pmatrix} \begin{pmatrix} 1 \\ x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}.$$

Cuando usamos coordenadas homogéneas deducidas de las referencias cartesianas $\mathcal{R}$ y $\mathcal{R}'$, es decir, si $(\xi^0, \xi^1, \ldots, \xi^n)$ con $x^i = \xi^i/\xi^0$ $(i = 1, \ldots, n)$ y si

$(\eta^0, \eta^1, \ldots, \eta^n)$ con $y^j = \eta^j/\eta^0$ $(j = 1, \ldots, m)$, las ecuaciones de la aplicación afín se expresa (poniendo $a_0^1 = b^1, \ldots, a_0^m = b^m$) por:

$$\begin{array}{l} \eta^0 = \xi^0 \\ \eta^1 = a_0^1\xi^0 + a_1^1\xi^1 + \cdots + a_n^1\xi^n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \eta^m = a_0^m\xi^0 + a_1^m\xi^1 + \cdots + a_n^m\xi^n \end{array} \quad \text{ó} \quad \begin{pmatrix} \eta^0 \\ \eta^1 \\ \vdots \\ \eta^m \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ a_0^1 & a_1^1 & \cdots & a_n^1 \\ \vdots & \vdots & \ddots & \vdots \\ a_0^m & a_1^m & \cdots & a_n^m \end{pmatrix} \begin{pmatrix} \xi^0 \\ \xi^1 \\ \vdots \\ \xi^n \end{pmatrix}.$$

Antes de tratar otros casos particulares de aplicaciones afines, a parte de las traslaciones, pongamos un par de ejemplos de aplicaciones afines en el plano real que tienen que ver con determinados fractales o con lenguajes de diseño gráfico; en las que interviene traslaciones, giros y homotecias en el plano afín real ([5, §3.3], http://webpages.ull.es/users/amontes/apuntes/gdh.pdf#h-afinidades).

4.19 Ejemplo.- *Operador* `makefont` *en PostScript como aplicación afín.*

En PostScript, las operaciones de traslación (`translate`), rotación (`rotate`) y cambio de escala (`scale`) quedan abarcadas en una transformación afín de la forma:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix} + \begin{pmatrix} m & s \\ r & n \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Con los seis coeficientes que intervienen en esta aplicación afín se confecciona la matriz del lenguaje PostScript siguiente [(1)]

$$[m\ r\ s\ n\ p\ q]$$

Uno de los usos de la matriz de transformación es para transformar caracteres de un texto por medio del operador `makefont`, o para modificar la representación de una imagen a partir de su descripción digitalizada, utilizando el operador `image`. Existen muchos más operadores que tienen como uno de sus argumentos una matriz de transformación (citemos, por ejemplo, aquellos que sustituyen una matriz por otra, el que multiplica matrices, el que obtiene la inversa, etc.).

En PostScript cada juego de caracteres está asociado a un matriz de transformación que determina la escala, la orientación y la posición de los caracteres. Esta matriz consta de seis elementos, con índices que varían del 0 al 5: $[m\ r\ s\ n\ p\ q]$

m es el parámetro (índice 0 de la matriz) que representa la longitud en puntos del carácter.

r (índice 1) está relacionado con el ángulo θ que forma el carácter con la horizontal, $r = m \operatorname{tag}\theta$.

s (índice 2) está relacionado con el ángulo β del carácter con la vertical, $s = n \operatorname{tag}\beta$.

n representa la altura en puntos del carácter.

p, q representan el desplazamiento en puntos, con respecto al origen de coordenadas; una posible utilización es para colocar subíndices y superíndices.

Esta matriz es el parámetro del operador `makefont` que transforma el juego de caracteres, por lo que $[m\ 0\ 0\ m\ 0\ 0]$ `makefont` es equivalente a m `scalefont`.

Podemos transformar el texto "Afinidad", mediante el siguiente programa:

---

(1) Más información en: http://webpages.ull.es/users/amontes/latexps/matriz.pdf.

```
/Times-Roman findfont [50 0 100 -100 0 0] makefont setfont
/texto (Afinidad)  def
20 -80 translate
0 0 moveto
gsave
   .5 setgray
   texto show
grestore
/Times-Roman findfont [50 0 50 50 0 0] makefont setfont
texto show
```

Cuyo resultado es el siguiente:

4.20 Ejemplo[Fractal punta de flecha].- *Con respecto a la referencia canónica en $\mathbb{R}^2$, considerar las tres transformaciones f, g y h dadas por*

$$f(x,y)=\Big(-\frac{1}{4}x-\frac{\sqrt{3}}{4}y+\frac{3}{4},\frac{\sqrt{3}}{4}x-\frac{1}{4}y+\frac{\sqrt{3}}{4}\Big),$$

$$g(x,y)=\Big(-\frac{1}{4}x+\frac{\sqrt{3}}{4}y-\frac{3}{4},-\frac{\sqrt{3}}{4}x-\frac{1}{4}y+\frac{\sqrt{3}}{4}\Big),$$

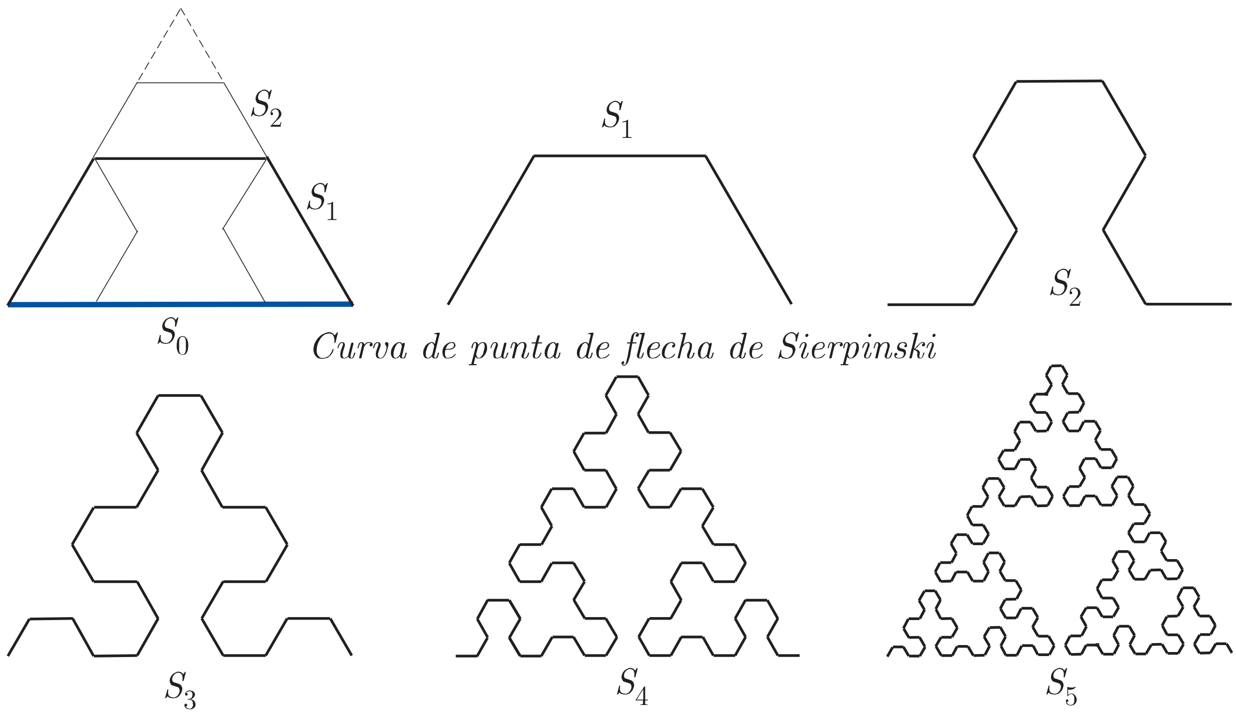
$$h(x,y)=\Big(\frac{1}{2}x,\frac{1}{2}y+\frac{\sqrt{3}}{4}\Big).$$

Las expresiones matriciales son, respectivamente:

$$\begin{pmatrix} x' \\ y' \end{pmatrix}=\begin{pmatrix} \frac{3}{4} \\ \frac{\sqrt{3}}{4} \end{pmatrix}+\frac{1}{2}\begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix}=\begin{pmatrix} -\frac{3}{4} \\ \frac{\sqrt{3}}{4} \end{pmatrix}-\frac{1}{2}\begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix}=\begin{pmatrix} 0 \\ \frac{\sqrt{3}}{4} \end{pmatrix}+\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$



Curva de punta de flecha de Sierpinski

(Applet-Cabri http://webpages.ull.es/users/amontes/cabri/g1-36.htm)

Estas tres aplicaciones afines resultan ser composición de giro, homotecia y traslación, como sigue:

f es la composición de un giro de centro en $O(0,0)$ y amplitud 120° con una homotecia de centro en O y razón $1/2$, seguida de una traslación de vector $\vec{\boldsymbol{v}} = (3/4, \sqrt{3}/4)$.

g es la composición de un giro de centro en $O(0,0)$ y amplitud 60° con una homotecia de centro en O y razón $-1/2$, seguida de una traslación de vector $\vec{\boldsymbol{w}} = (-3/4, \sqrt{3}/4)$.

h es la composición de homotecia de centro $O(0,0)$ y razón $1/2$ con la traslación de vector $\vec{\boldsymbol{u}} = (0, \sqrt{3}/4)$. O bien, h es una homotecia de centro $(0, \sqrt{3}(2)$ y razón $1/2$.

El segmento $S_0 = [-1,1]$ se trasforma en la poligonal $S_1 = f(S_0) \cup g(S_0) \cup h(S_0)$ (asociada al triángulo equilátero de base S_0), que parte del punto $(1,0)$ hasta la mitad del lado derecho (ésta es la imagen $f(S_0)$), seguida del segmento que une los puntos medios de los lados laterales ($h(S_0)$) y de la mitad del lado izquierdo hasta llegar a $(-1,0)$ ($g(S_0)$).

Procediendo iteradamente se obtiene una sucesión de líneas quebradas

$$S_{n+1} = f(S_n) \cup g(S_n) \cup h(S_n),$$

que forman el fractal punta de flecha de Sierpinski [(2)].

(2) Waclaw Sierpinski (Varsovia, 1882-1969) Matemático polaco. Miembro fundador de la escuela matemática polaca moderna, junto con Janiszewski y Mazurkiewicz, que contribuyó al progreso de la teoría de conjuntos y de la topología y favoreció la consolidación de los fundamentos lógicos de las matemáticas. Llevó a cabo importantes investigaciones sobre teoría de números. Dedicó una parte de sus investigaciones al estudio de distintas formas de fractales, como ejemplos: (http://www.cut-the-knot.org/Curriculum/Geometry/SierpinskiTremas.shtml)

4.4 Homotecias

4.21 Definición.- *En un espacio afín $(\mathcal{A}, \mathbf{E}, +)$, fijado un punto P y un escalar $k \in K - \{0, 1\}$, una* ***homotecia*** *de centro P y razón k es la aplicación $h_{P,k}\colon \mathcal{A} \to \mathcal{A}$ definida por $h_{P,k}(X) = P + k\overrightarrow{PX}$, para todo $X \in \mathcal{A}$.*

4.22 Proposición.- *Una homotecia es una aplicación afín de aplicación lineal asociada $\tilde{h}_{P,k} = k\,1_{\mathbf{E}}$. Una homotecia $h_{P,k}$ es una afinidad de aplicación inversa la homotecia $h_{P,\frac{1}{k}}$.*

Demostración.- Sean $S = h_{P,k}(Q) = P + k\,\overrightarrow{PQ}$ y $T = h_{P,k}(R) = P + k\,\overrightarrow{PR}$, es decir, $\overrightarrow{PS} = k\,\overrightarrow{PQ}$ y $\overrightarrow{PT} = k\,\overrightarrow{PR}$. Luego, $\overrightarrow{ST} = \overrightarrow{PT} - \overrightarrow{PS} = k(\overrightarrow{PR} - \overrightarrow{PQ}) = k\,\overrightarrow{QR}$. Así, la aplicación lineal asociada $\tilde{h}_{P,k}$ verifica:

$$\tilde{h}_{P,k}(\overrightarrow{QR}) = \overrightarrow{h_{P,k}(Q)h_{P,k}(R)} = \overrightarrow{ST} = k\,\overrightarrow{QR}.$$

Se tiene que $\tilde{h}_{P,k}(\vec{\boldsymbol{v}}) = k\vec{\boldsymbol{v}}$, que es claramente lineal y biyectiva, con aplicación inversa $\tilde{h}^{-1}_{P,k}(\vec{\boldsymbol{v}}) = \frac{1}{k}\vec{\boldsymbol{v}}$. ◪

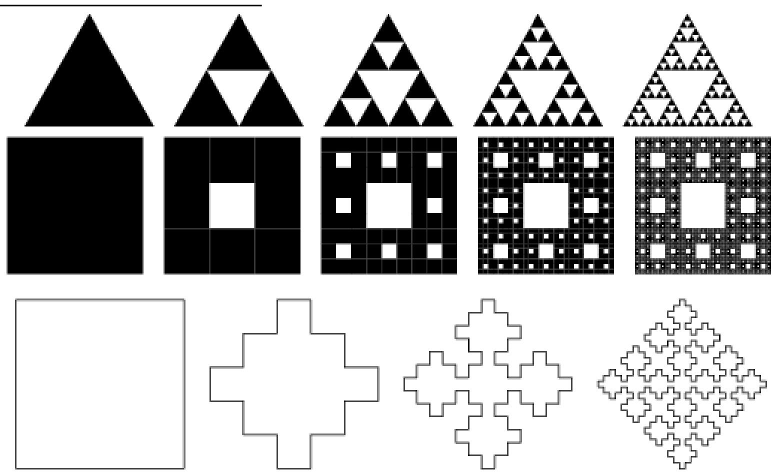
4.23 Proposición.- *Una recta y su imagen, mediante una homotecia, son paralelas.*

Demostración.- Sean una homotecia $h_{P,k}$ y una recta $\mathcal{L}$ que contiene al punto Q y tiene la dirección de un vector $\vec{\boldsymbol{v}}$. Un punto genérico de $\mathcal{L}$ es de la forma $Q + \lambda\vec{\boldsymbol{v}}$ $(\lambda \in K)$, entonces su imagen mediante la homotecia $h_{P,k}$ es:

$$h_{P,k}(Q + \lambda\vec{\boldsymbol{v}}) = P + k\overrightarrow{P(Q+\lambda\vec{\boldsymbol{v}})} = P + k\overrightarrow{PQ} + k\lambda\vec{\boldsymbol{v}} = h_{P,k}(Q) + \lambda'\vec{\boldsymbol{v}},$$

que está en la recta $\mathcal{L}$' que pasa por $h_{P,k}(Q)$ y tiene la dirección del vector $\vec{\boldsymbol{v}}$, es decir, paralela a $\mathcal{L}$. ◪

Una homotecia $h_{P,k}$ tiene a P como único punto fijo y hemos visto cómo es la aplicación lineal asociada a una homotecia; a modo de recíproco, se tiene



que, si una aplicación afín tiene como aplicación lineal asociada la aplicación identidad multiplicada por un escalar, ella es una homotecia, de la cual vamos a determinar su punto fijo.

Para poder encontrar la aplicación afín $f\colon \mathcal{A} \to \mathcal{A}$ que tenga como aplicación lineal asociada $k1_E$, necesitamos conocer la imagen R de un punto Q (pág 59). Se ha de verificar que $f(X) = f(Q) + k1_E(\overrightarrow{QX}) = R + k\overrightarrow{QX}$. Si f ha de ser una homotecia ha de tener un punto fijo P, que ha de cumplir $P = R + k\overrightarrow{QP}$, o sea, $\overrightarrow{RP} = \overrightarrow{RQ} + \overrightarrow{QP} = k\overrightarrow{QP}$; entonces, el punto fijo verifica:

$$\overrightarrow{QP} = \frac{1}{1-k}\overrightarrow{QR} \Leftrightarrow P = Q + \frac{1}{1-k}\overrightarrow{QR}.$$

La homotecia $h_{P,k}$ coincide f, ya que

$$\begin{aligned} h_{P,k}(X) &= P + k\overrightarrow{PX} = Q + \frac{1}{1-k}\overrightarrow{QR} + k(\overrightarrow{PQ} + \overrightarrow{QX}) = \\ &= Q + \frac{1}{1-k}\overrightarrow{QR} - k(\frac{1}{1-k}\overrightarrow{QR}) + k\overrightarrow{QX} = \\ &= Q + \overrightarrow{QR} + k\overrightarrow{QX} = R + k\overrightarrow{QX} = f(X). \end{aligned}$$

4.24 Nota.- Si de una homotecia se conocen su razón k y la imagen R de un punto Q, su centro P es baricentro de puntos ponderados $(Q, -k)$ y $(R, 1)$. O lo que es lo mismo, P es el punto tal que la razón simple $(P\,Q\,R) = k$.

Ya que su centro, como se acaba de exponer, es el punto

$$\overrightarrow{QP} = \frac{1}{1-k}\overrightarrow{QR} \Leftrightarrow (1-k)\overrightarrow{QP} = \overrightarrow{QR} \Leftrightarrow -k\overrightarrow{PQ} + \overrightarrow{PR} = \vec{\mathbf{0}}.$$

La última igualdad expresa, si $1 - k \neq 0$, que P es el baricentro de Q y R afectados de los "pesos" $-k$ y 1, respectivamente (Definición de baricentro, pág. 40).

4.25 Ejemplo.- *Un caso particular de homotecias son las que poseen razón $k = -1$, $h_{P,-1} = s_P$, a la que se le denomina **simetría central** de centro P; esto es, $s_P(X) = P - \overrightarrow{PX}$ o sea $\overrightarrow{Ps_P(X)} + \overrightarrow{PX} = \vec{\mathbf{0}}$, es decir P es el punto medio de X y $s_P(X)$.*

Si componemos traslaciones y homotecias resultarán afinidades del mismos tipo, por lo que el conjunto de traslaciones y homotecias, con la operación composición, es un grupo, subgrupo de grupo afín.

Veamos cuál es el centro y razón de la homotecia resultante o el vector de la traslación, en su caso.

La aplicación lineal asociada a la composición de dos homotecias, h_{P,k_1} y h_{Q,k_2}, es la composición de las aplicaciones lineales asociadas a cada una de ellas, o sea, $\vec{\boldsymbol{v}} \mapsto k_1k_2\vec{\boldsymbol{v}}$. Por tanto, la composición de las dos homotecias es una homotecia de razón k_1k_2.

Si su centro es el punto R, para todo $X \in \mathcal{A}$, se tiene que $\overrightarrow{RX"} = k\overrightarrow{RX}$, siendo $X" = h_{Q,k_2}(X')$ y $X' = h_{P,k_1}(X)$.

$$\overrightarrow{RX"} = \overrightarrow{RQ} + \overrightarrow{QX"} = \overrightarrow{RQ} + k_2\overrightarrow{QX'} =$$
$$= \overrightarrow{RQ} + k_2\overrightarrow{QP} + k_1k_2\overrightarrow{PX} = \overrightarrow{RQ} + k_2\overrightarrow{QP} + k_1k_2\overrightarrow{PR} + k_1k_2\overrightarrow{RX}.$$

De aquí se sigue que

$$k = k_1k_2, \qquad \overrightarrow{RQ} + k_2\overrightarrow{QP} + k_1k_2\overrightarrow{PR} = \vec{\boldsymbol{0}}.$$
$$\overrightarrow{RP} + \overrightarrow{PQ} + k_2\overrightarrow{QP} + k_1k_2\overrightarrow{PR} = \vec{\boldsymbol{0}}.$$
$$\overrightarrow{PR} - k_1k_2\overrightarrow{PR} = \overrightarrow{PQ} - k_2\overrightarrow{PQ}.$$
$$R = P + \frac{1-k_2}{1-k_1k_2}\overrightarrow{PQ}.$$

Por tanto, la composición de las homotecias h_{P,k_1} y h_{Q,k_2} es la homotecia de centro alineado con P y Q, en $P + \dfrac{1-k_2}{1-k_1k_2}\overrightarrow{PQ}$ y razón k_1k_2, el producto de razones.

Otra manera de determinar el centro R de la homotecia $h_{Q,k_2} \circ h_{P,k_1}$, que tiene razón k_1k_2, es usar el hecho de que R es el baricentro de un punto y su imagen, afectados de los "pesos" $-k_1k_2$ y 1, respectivamente. Así, sea P y su imagen $P" = (h_{Q,k_2} \circ h_{P,k_1})(P) = Q + k_2\overrightarrow{QP}$, entonces

$$-k_1k_2\overrightarrow{RP} + \overrightarrow{RP"} = \vec{\boldsymbol{0}} \Leftrightarrow -k_1k_2\overrightarrow{RP} + \overrightarrow{RQ} + k_2\overrightarrow{QP} = \vec{\boldsymbol{0}} \Leftrightarrow$$

$$-k_1k_2\overrightarrow{RP} + \overrightarrow{RP} + \overrightarrow{PQ} + k_2\overrightarrow{QP} = \vec{\boldsymbol{0}} \Leftrightarrow -(1-k_1k_2)\overrightarrow{PR} + (1-k_2)\overrightarrow{PQ} = \vec{\boldsymbol{0}} \Leftrightarrow$$

$$R = P + \frac{1-k_2}{1-k_1k_2}\overrightarrow{PQ}.$$

El centro de la homotecia composición obtenido tiene sentido si $k_1k_2 \neq 1$.

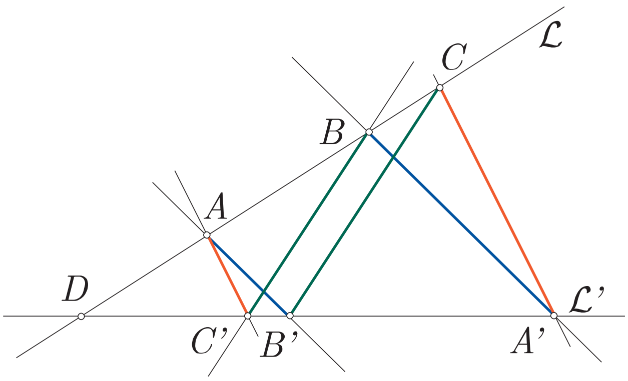
Cuando $k_1k_2 = 1$ la aplicación lineal asociada a la composición es la identidad, por tanto, se trata de una traslación si $k_1 \neq 1$ ($k_2 \neq 1$) o de la identidad si $k_1 = k_2 = 1$ (cada homotecia de partida es la identidad).

Como aplicación de las traslaciones y homotecias tratamos, a continuación, unas versiones afines de los teoremas clásicos de Pappus y Desargues. En realidad, estos resultados son teoremas de la geometría proyectiva, y de tal punto de vista se pueden demostrar estos teorema en una versión más general ([6, pág. 130 y 135]).

4.26 Proposición[Teorema de Pappus].- *Dados dos rectas $\mathfrak{L}$ y $\mathfrak{L}'$ sobre un plano afín. Sean A, B y C puntos sobre la primera recta, y A', B' y C' puntos sobre la segunda recta tales que ninguno de ellos coincide con el punto D de intersección de $\mathfrak{L}$ y $\mathfrak{L}'$ (si se se intersecan), si la recta AB' es paralela a la BA', y la recta BC' es paralela a CB', entonces las rectas AC' y CA' son paralelas.*

Demostración.- Si las rectas $\mathfrak{L}$ y $\mathfrak{L}'$ no son paralelas, sea D su punto de intersección.

(Applet-Cabri http://webpages.ull.es/users/amontes/cabri/g1-34.htm)



Sean f y g las homotecias de mismo centro D y tal que $f(A) = B$ y $g(B) = C$. Usando la Proposición 4.23, se tiene que $A' = f(B')$ y $B' = g(C')$. Además, como las homotecias de mismo centro conmutan, $h = f \circ g = g \circ f$ es una homotecia de centro D, se obtiene que $C = h(A)$ y $A' = h(C')$. Con lo que las rectas AC' y CA' son paralelas.

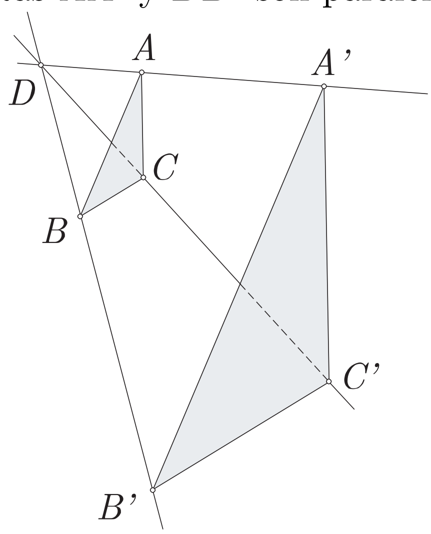
Si las rectas $\mathfrak{L}$ y $\mathfrak{L}'$ son paralelas se usan traslaciones en vez de homotecias. ◪

4.27 Proposición[Teorema de Desargues].- *En un espacio afín $(\mathcal{A}, \mathbf{E})$ se dan dos triángulos $\widehat{ABC}$ y $\widehat{A'B'C'}$, donde los vértices A, B, C, A', B', C' son todos distintos, si las rectas AB y $A'B'$ son paralelas y también las rectas BC y $B'C'$ son paralelas, entonces:*

Las rectas BC y $B'C'$ son paralelas $\Leftrightarrow$

Las rectas AA', BB', CC' son paralelas o concurrentes.

Demostración.- ($\Rightarrow$) Ya que las rectas AB y $A'B'$ son paralelas, los puntos A, A', B y B' son coplanarios ($\overrightarrow{AB'} = \overrightarrow{AA'} + \overrightarrow{A'B'} = \overrightarrow{AA'} + \lambda\overrightarrow{AB}$). Así, las rectas AA' y BB' son paralelas o concurrentes.



Consideremos el segundo caso, cuando ellas se intersecan, y sea D su punto intersección. Sean f la homotecia de centro D y tal que $f(A) = A'$. Por la Proposición 4.23, $f(B) = B'$.

Si $f(C) = C''$, la recta AC es paralela a $A'C''$ y, por tanto, las rectas $A'C'$ y $A'C''$ coinciden. Análogamente, se tiene que la recta BC es paralela a $B'C''$ y, por tanto, las rectas $B'C'$ y $B'C''$ coinciden. En consecuencia, $C' = C''$. Luego, C y $C' = f(C)$ están alineados con el centro de homotecias D.

La demostración en el caso de que las rectas AA' y BB' sean paralelas, se deja como ejercicio.

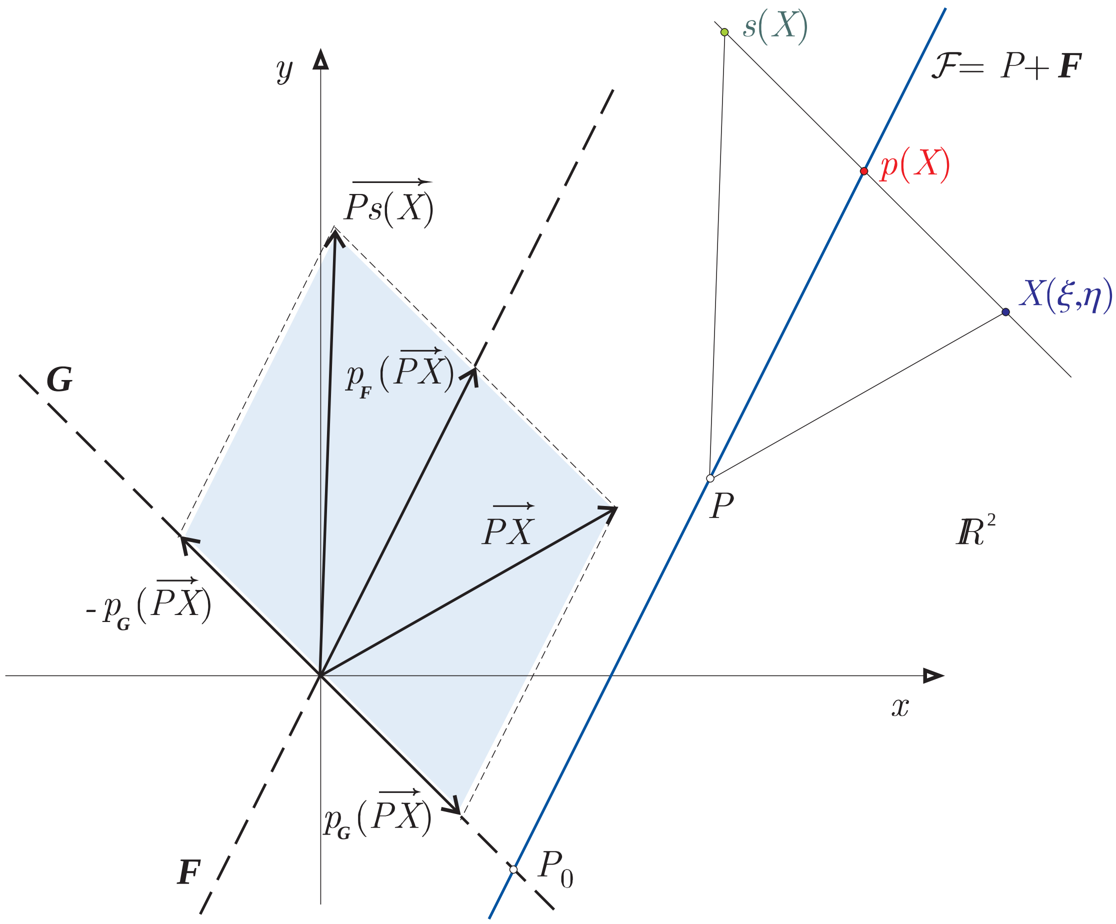
($\Leftarrow$) Es similar a la demostración en el otro sentido y se deja como ejercicio. ◪

4.5 Proyecciones

Antes de dar la definición de proyecciones en el lenguaje de aplicaciones afines, vemos unos ejemplos en los espacios vectoriales $\mathbb{R}^2$ y $\mathbb{R}^3$, con las estructuras de espacio afín canónicas.

4.28 Ejemplo.- *Sean $\boldsymbol{F} = \{(x,y) \in \mathbb{R}^2 / y = 2x\}$ y $\boldsymbol{G} = \{(x,-x)/x \in \mathbb{R}\}$ dos subespacios vectoriales suplementarios de $\mathbb{R}^2$. Consideramos la variedad lineal (recta) $\mathcal{F} = P + \boldsymbol{F}$ que pasa por $P = (2,1)$ y tiene la dirección de $\boldsymbol{F}$. Dado un punto $X \in \mathbb{R}^2$, su proyección sobre $\mathcal{F}$ en la dirección de $\boldsymbol{G}$ (es decir, en la dirección de cualquier variedad lineal $\mathcal{G}$ de subespacio director $\boldsymbol{G}$) es el punto $p(X)$ obtenido considerando la componente única del vector $\overrightarrow{PX}$ en $\boldsymbol{F}$ en la descomposición de $\mathbb{R}^2$ en suma directa $\mathbb{R}^2 = \boldsymbol{F} \oplus \boldsymbol{G}$, que llamamos $p_{\boldsymbol{F}}(\overrightarrow{PX})$; con lo que podemos poner:*

$$p(X) = P + \tilde{p}(\overrightarrow{PX}).$$



(Observación: Al punto $s(X)$ que aparece en esta figura no nos referimos aquí, sino en párrafo 4.6, dedicado a simetrías.)

La expresión en coordenadas de esta proyección se puede obtener por dos vías: una, la que se acaba de exponer y, otra, mediante intersección de rectas. Así:

1) Sea un punto $X(\xi,\eta)$. Obtención de la proyección descomponiendo el vector $\overrightarrow{PX}$ en $\boldsymbol{F} \oplus \boldsymbol{G}$, tomando como base $\{(1,2)\}$ en $\boldsymbol{F}$ y $\{(-1,1)\}$ en $\boldsymbol{G}$:

$$(\xi-2,\eta-1) = \lambda(1,2) + \mu(-1,1) \Rightarrow \lambda = \frac{1}{3}(\xi+\eta-3), \quad \mu = \frac{1}{3}(-2\xi+2\eta+3).$$

$$p(\xi,\eta) = (2,1) + \frac{1}{3}(\xi+\eta-3)(1,2); \qquad \begin{pmatrix} \xi' \\ \eta' \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}.$$

Antes de exponer el segundo método para determinar p, indiquemos que el resultado obtenido no depende del punto P tomado en $\mathcal{F}$, como se comprueba sustituyéndolo por un punto genérico $Q(t, 2t-3)$. Las ecuaciones de la proyección obtenida tampoco dependen de las bases tomadas en los subespacios $\boldsymbol{F}$ y $\boldsymbol{G}$.

La proyección sobre $\mathcal{F}$ obtenida es una aplicación afín (ver Ejemplo 4.3): composición de un endomorfismo en $I\!R^2$ (que en este caso es un proyector –Definición 1.39– en $I\!R^2$) seguida de una traslación de vector $\vec{\boldsymbol{v}} = (1, -1)$, cuyo extremo está en $\mathcal{F}$

Tomando $P_0(1, -1)$, se puede escribir $p(X) = p(P_0) + p_{\boldsymbol{F}}(\overrightarrow{P_0X})$; con lo que p es una aplicación afín de aplicación lineal asociada el proyector $p_{\boldsymbol{F}}$

2) Sea un punto $X(\xi, \eta)$, para determinar su proyección $p(X)$, se intersecan la recta que pasa por $X(\xi, \eta)$ y es paralela a $\boldsymbol{G}$ con la recta que pasa por $P(2, 1)$ y es paralela a $\boldsymbol{F}$.

$$-(x-\xi) = y - \eta, \qquad y - 1 = 2(x-2).$$

4.29 Ejemplo.- *Sean $\boldsymbol{F} = \{(x, y, 0) \in I\!R^3 / x, y \in I\!R\}$ y $\boldsymbol{G} = \{(0, t, t)/t \in I\!R\}$ dos subespacios vectoriales complementarios de $I\!R^3$. Consideramos la variedad lineal (plano) $\mathcal{F} = P + \boldsymbol{F}$ que pasa por $P = (0, 0, 1)$ y tiene la dirección de $\boldsymbol{F}$. Dado un punto $X \in I\!R^3$, su proyección sobre $\mathcal{F}$ en la dirección de $\boldsymbol{G}$ (es decir, en la dirección de cualquier variedad lineal $\mathcal{G}$ de subespacio director $\boldsymbol{G}$) es el punto $p(X)$ obtenido considerando la componente única del vector $\overrightarrow{PX}$ en $\boldsymbol{F}$ en la descomposición de $I\!R^3$ en suma directa $I\!R^3 = \boldsymbol{F} \oplus \boldsymbol{G}$, que llamamos $p_{\boldsymbol{F}}(\overrightarrow{PX})$.*

Si $X = (\xi, \eta, \zeta) \in I\!R^3$, la proyección $p(X) = (\xi', \eta', \zeta')$ de X sobre el plano $\mathcal{F}$ paralelamente a la recta $\boldsymbol{G}$ se obtiene intersecando el plano $z = 1$ con la recta que pasa por (ξ, η, ζ) y paralela a la recta $x = 0, y = z$, cuyas ecuaciones son $x - \xi = 0, y - \eta = z - \zeta$; resulta:

$$(\xi', \eta', \zeta') = (\xi, \eta - \zeta + 1, 1).$$

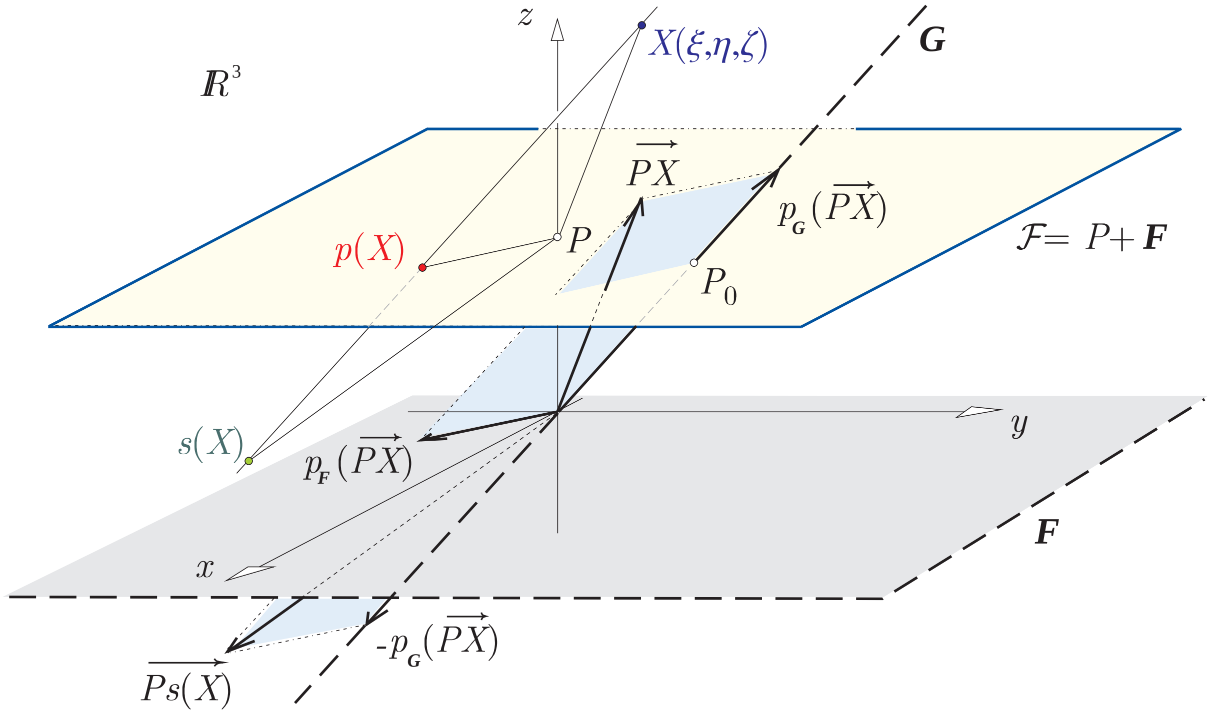
O bien, resolviendo en λ, μ y ν:

$$(\xi, \eta, \zeta - 1) = \lambda(1, 0, 0) + \mu(0, 1, 0) + \nu(0, 1, 1) \Rightarrow \lambda = \xi, \mu = \eta - \zeta + 1, \nu = \zeta - 1$$

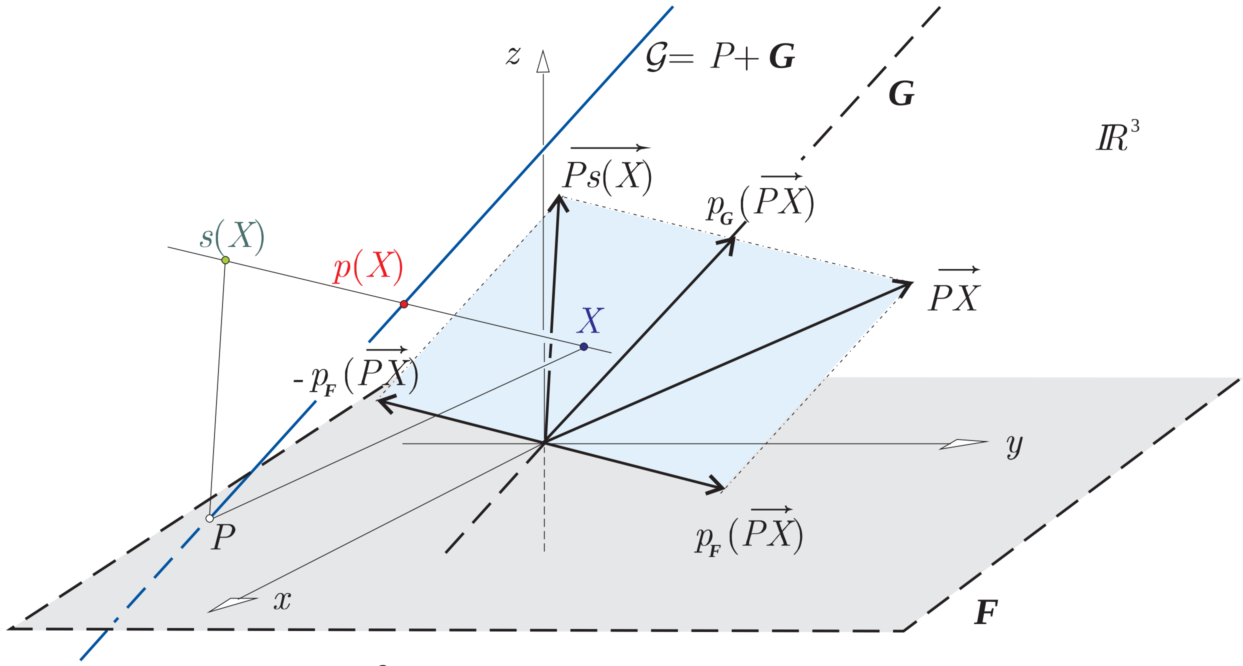
$$(\xi', \eta', \zeta') = (0, 0, 1) + \xi(1, 0, 0) + (\eta - \zeta + 1)(0, 1, 0),$$

$$\begin{pmatrix} \xi' \\ \eta' \\ \zeta \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}.$$

Así, si $P_0 = (0, 1, 1)$, $p(X) = p(P_0) + p_{\boldsymbol{F}}(\overrightarrow{P_0X})$ y p es una aplicación afín.



4.30 Ejemplo.- *Sean $\mathbf{F} = \{(x, y, 0) \in I\!R^3 / x, y \in I\!R\}$ y $\mathbf{G} = \{(0, t, t)/t \in I\!R\}$ dos subespacios vectoriales suplementarios de $I\!R^3$. Consideramos la variedad lineal (recta) $\mathcal{G} = P + \mathbf{G}$ que pasa por $P = (1, -1, 0)$ y tiene la dirección de $\mathbf{G}$. Dado un punto $X \in I\!R^3$, su proyección sobre $\mathcal{G}$ en la dirección de $\mathbf{F}$ (es decir, en la dirección de cualquier variedad lineal $\mathcal{F}$ de subespacio director $\mathbf{F}$) es el punto $p(X)$ obtenido considerando la componente única del vector $\overrightarrow{PX}$ en $\mathbf{G}$ en la descomposición de $I\!R^3$ en suma directa $I\!R^3 = \mathbf{F} \oplus \mathbf{G}$, que llamamos $p_{\mathbf{F}}(\overrightarrow{PX})$.*



Si $X = (\xi, \eta, \zeta) \in I\!R^3$, la proyección $p(X) = (\xi', \eta', \zeta')$ de X sobre la recta $\mathcal{G}$ paralelamente al plano $\mathbf{F}$ se obtiene intersecando la recta $z = 1, y + 1 = z$ con el plano que pasa por (ξ, η, ζ) y paralelo al $z = 0$; resulta:

$$(\xi', \eta', \zeta') = (1, \zeta - 1, \zeta).$$

O bien, resolviendo en λ, μ y ν:
$(\xi-1, \eta+1, \zeta) = \lambda(1,0,0)+\mu(0,1,0)+\nu(0,1,1) \Rightarrow \lambda = \xi-1, \mu = \eta-\zeta+1, \nu = \zeta$.

$$(\xi', \eta', \zeta') = (1,-1,0) + \zeta(0,1,1),$$

$$\begin{pmatrix} \xi' \\ \eta' \\ \zeta \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}.$$

Así, $p(X) = p(P) + p_F(\overrightarrow{PX})$ y p es una aplicación afín.

En cada uno de estos ejemplos la proyección p, sobre una variedad afín (recta o plano) paralelamente a otra variedad afín (recta o plano), es una aplicación afín, de aplicación lineal asociada el proyector p_F ($p_F^2 = p_F$). Pues bien, este puede ser el punto de partida para definir una proyección en un espacio afín $\mathcal{A}$, como una aplicación afín de aplicación lineal asociada un proyector conocido, dando además, para poder determinar la aplicación afín, la imagen de un punto u otra condición que sustituya a ésta.

4.31 Definición.- *Sea $\mathcal{A}$ un espacio afín asociados a un espacio vectorial $\boldsymbol{E}$; dadas dos variedades lineales $\mathcal{F}$ y $\mathcal{G}$ con subespacios directores $\boldsymbol{F}$ y $\boldsymbol{G}$, respectivamente, tal que $\boldsymbol{F} \oplus \boldsymbol{G} = \boldsymbol{E}$ ($\boldsymbol{F}$ y $\boldsymbol{G}$ son suplementarios), consideramos el proyector p_F de $\boldsymbol{E}$ sobre $\boldsymbol{F}$ (paralelamente a $\boldsymbol{G}$) y un punto $P \in \mathcal{F}$, a la aplicación (afín) $p_{\mathcal{F}}: \mathcal{A} \to \mathcal{A}$ definida, para todo X, por*

$$p_{\mathcal{F}}(X) = P + p_F(\overrightarrow{PX})$$

*le llamamos **proyección** sobre $\mathcal{F}$ paralelamente a $\mathcal{G}$.*

Para que esta definición sea coherente, debemos confirmar que la aplicación $p_{\mathcal{F}}: \mathcal{A} \to \mathcal{A}$ es verdaderamente afín y, además, que no depende del punto P particular, tomado sobre $\mathcal{F}$.

En efecto, si en la definición de $p_{\mathcal{F}}$ tomamos $X = P$, se tiene que $p_{\mathcal{F}}(P) = P$ y, por tanto,

$$p_{\mathcal{F}}(X) = p_{\mathcal{F}}(P) + p_F(\overrightarrow{PX}).$$

Así, $p_{\mathcal{F}}$ es una aplicación afín de aplicación lineal asociada $\widetilde{p_{\mathcal{F}}}$ el proyector p_F.

Veamos que el punto P no es relevante en la definición. Sea otro punto $Q \in \mathcal{F}$, entonces $\overrightarrow{PQ} \in \boldsymbol{F}$ y $p_F(\overrightarrow{PQ}) = \overrightarrow{PQ}$; se tiene que

$$Q + p_F(\overrightarrow{QX}) = P + p_F(\overrightarrow{PX}).$$

Pues, $Q + p_F(\overrightarrow{QX}) = Q + p_F(\overrightarrow{QP} + \overrightarrow{PX}) = Q + \overrightarrow{QP} + p_F(\overrightarrow{PX}) = P + p_F(\overrightarrow{PX})$.

A continuación enunciamos unas primeras propiedades de una proyección $p_{\mathcal{F}}$ sobre una variedad lineal $\mathcal{F}$ (de subespacio vectorial asociado $\boldsymbol{F}$) paralelamente a $\mathcal{G}$ (de subespacio vectorial asociado $\boldsymbol{G}$). Todas ellas se deducen de forma inmediata a partir de la definición de proyección que hemos dado.

1) $\mathcal{F} = \mathrm{Im}(p_{\mathcal{F}}) = f(\mathcal{A})$ es el conjunto de punto fijos.

2) $\forall X \in \mathcal{A} \Rightarrow \{p_{\mathcal{F}}(X)\} = \mathcal{F} \cap (X + \boldsymbol{G})$.

3) $\forall P \in \mathcal{F} \Rightarrow (X \in P + \boldsymbol{G} \Leftrightarrow p_F(X) = P)$.

4) $p_{\mathcal{F}}^2 = p_{\mathcal{F}} \circ p_{\mathcal{F}} = p_{\mathcal{F}}$.

5) $\forall X \in \mathcal{A} \Rightarrow p_{F}(\overrightarrow{Xp_{\mathcal{F}}(X)}) = \vec{0}_E$.

La propiedad 4) caracteriza las aplicaciones afines que son proyecciones:

4.31' Definición.- *Una aplicación afín $f\colon(\mathcal{A}, E) \to (\mathcal{A}, E)$, que verifica $f^2 = f \circ f = f$, se dice que es una **proyección**.*

Las dos definiciones que hemos dados son equivalentes y para establecerlo nos basta, a partir de esta última, determinar las variedades lineales $\mathcal{F}$ y $\mathcal{G}$ que intervienen en la Definición 4.31.

Observemos primero que la aplicación lineal $\tilde{f}$ asociada es a f también verifica $\tilde{f}^2 = \tilde{f}$. Ya que, para dos puntos arbitrarios $P, Q \in \mathcal{A}$:

$$\tilde{f}^2(\overrightarrow{PQ}) = \tilde{f}(\overrightarrow{f(P)f(Q)}) = \overrightarrow{f^2(P)f^2(Q)} = \overrightarrow{f(P)f(Q)} = \tilde{f}(\overrightarrow{PQ}).$$

Esto quiere decir que $\tilde{f}$ es el proyector $p_{\tilde{f}(E)}$ de E sobre $\tilde{f}(E)$ (paralelamente a $\mathrm{Ker}(\tilde{f})$).

Por otra parte, $\mathcal{F} = f(\mathcal{A})$ queda fijo por f; pues si $Y \in f(\mathcal{F})$, existe un punto $X \in \mathcal{A}$ tal que $Y = f(X)$, entonces $f(Y) = f^2(X) = f(X) \in f(\mathcal{A})$.

Tomamos la variedad lineal (Proposición 4.4) $\mathcal{F} = f(\mathcal{A}) = \mathrm{Im}(f)$, cuyo subespacio director es $F = \tilde{f}(E)$ y una variedad lineal tal que su subespacio director sea $G = \mathrm{Ker}(\tilde{f})$. Entonces, f es la proyección sobre $\mathcal{F}$ paralelamente a cualquier variedad lineal de subespacio director G, ya que, fijado un punto $P \in f(\mathcal{A})$, para todo $X \in \mathcal{A}$, se tiene que

$$f(X) = P + p_{\tilde{f}(E)}(\overrightarrow{PX}).$$

4.32 Ejemplo.- *Consideremos la aplicación $f\colon \mathbb{R}^3 \to \mathbb{R}^3$, definida por*

$$f(x, y, z) = \left(\frac{x+y}{2} + 1, \frac{x+y}{2} - 1, 0\right), \text{ ó } X' = P + MX:$$

$$\begin{pmatrix} x' \\ y' \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

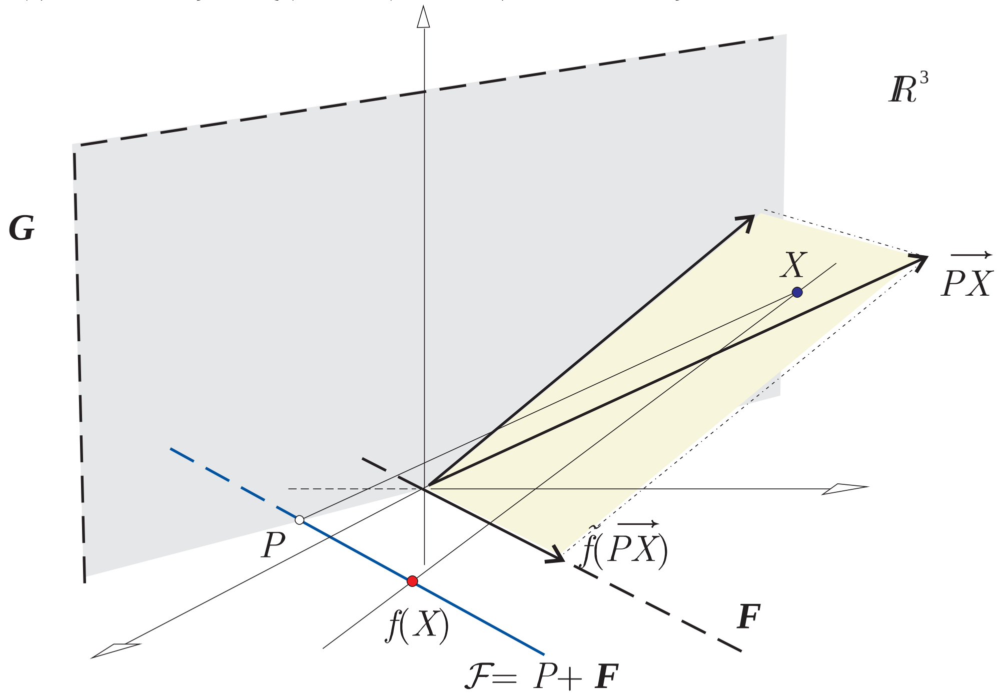
Se trata de una proyección sobre una recta que pasa por $P(1, -1, 0)$ paralela al plano $x + y = 0$.

Como $\big((x+y)/2 + 1 + (x+y)/2 - 1\big)/2 = (x+y)/2$, se ve que $f^2 = f$: se trata de una proyección en el espacio afín $\mathbb{R}^3$. Su aplicación lineal asociada $\tilde{f}$ tiene como matriz M, respecto a la base canónica en el espacio vectorial $\mathbb{R}^3$

Todo punto se proyecta en una recta, pues la dimensión de $\tilde{f}(\mathbb{R}^3)$ es 1, que coincide con el rango de la matriz M. Un vector que da la dirección de la recta es $(1/2, 1/2, 0)$. Se trata pues de la recta que pasa por el punto $P(1, -1, 0)$ y dirección $(1, 1, 0)$:

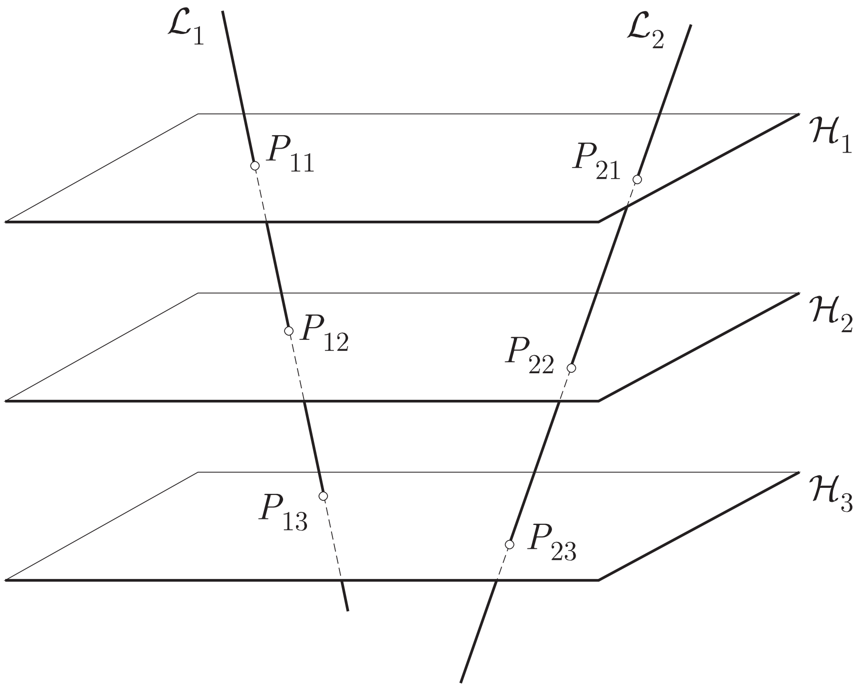
$$\mathcal{F} = \{(x, y, z) \in \mathbb{R}^3 / x - y - 2 = 0, z = 0\}.$$

La dirección de la proyección es la del subespacio vectorial $\boldsymbol{G} = \text{Ker}(f) = \{(x, -x, z)/x, z \in \mathbb{R}\} = \{(x, y, z) \in \mathbb{R}^3/x + y = 0\}$.



Como aplicación de las proyecciones, damos otra demostración del teorema de Tales, distinta de la que hemos expuesto en la Proposición 3.27.

4.33 Proposición[Teorema de Tales].- *En un espacio afín $(\mathcal{A}, \boldsymbol{E})$, dados tres hiperplanos paralelos $\mathcal{H}_1, \mathcal{H}_2$ y $\mathcal{H}_3$ y dos rectas, no paralelas a los hiperplanos, $\mathcal{L}_1$ y $\mathcal{L}_2$ tales que $\mathcal{L}_i \cap \mathcal{H}_j = \{P_{ij}\}$, entonces se tiene la igualdad de las razones simples $(P_{11}\, P_{12}\, P_{13}) = (P_{21}\, P_{22}\, P_{23})$.*



Demostración.- Consideremos, para $i = 1, 2$, las proyecciones $p_i: \mathcal{A} \to \mathcal{A}$ sobre la recta $\mathcal{L}_i$ paralelamente a los hiperplanos dados de subespacio director $\boldsymbol{H}$. Estas proyecciones existen pues rectas e hiperplanos contiene un sólo punto,

por los que sus subespacios directores son suplementarios. Entonces, por la propiedad 2) de la página 80,

$$p_2(P_{1j}) = \mathcal{L}_2 \cap (P_{ij} + \boldsymbol{H}) = \mathcal{L}_2 \cap \mathcal{H}_j = P_{2j}.$$

Ahora, sea $\rho = (P_{11}\, P_{12}\, P_{13})$, es decir, $\overrightarrow{P_{11}P_{13}} = \rho\overrightarrow{P_{11}P_{12}}$; entonces,

$$\overrightarrow{P_{21}P_{23}} = \overrightarrow{p_2(P_{11})p_2(P_{13})} = \widetilde{p_2}(\overrightarrow{P_{11}P_{13}}) = \widetilde{p_2}(\lambda\,\overrightarrow{P_{11}P_{12}}) = \lambda\,\widetilde{p_2}(\overrightarrow{P_{11}P_{12}}) =$$

$$= \lambda\,\overrightarrow{p_2(P_{11})p_2(P_{12})} = \lambda\overrightarrow{P_{21}P_{22}}.$$

Con lo que $\lambda = (P_{21}\, P_{22}\, P_{23})$.

4.6 Simetrías

Volvemos a los ejemplos previos a la definición de proyección. Así, en el Ejemplo 4.28, consideramos, para cada punto X en $\mathbb{R}^2$, el punto $s(X)$ tal que $p(X)$ (proyección de X sobre la recta $\mathcal{F}$ en la dirección de una recta $\boldsymbol{G}$) es el punto medio de los puntos X y $s(X)$. Es decir, $s(X)$ es el simétrico de X respecto a $\mathcal{F}$ en la dirección de $\boldsymbol{G}$. Se tiene, entonces, que $\overrightarrow{Xs(X)} = 2\overrightarrow{Xp(X)}$, o lo que es lo mismo

$$s(X) = X + 2\overrightarrow{Xp(X)}.$$

Además, si $P \in \mathcal{F}$ se cumple que $s(P) = P$; en particular, podemos tomar $P(2,1)$, como en dicho ejemplo. Así, el punto $s(X)$ lo podemos obtener, utilizando los proyectores $p_{\boldsymbol{F}}$ y $p_{\boldsymbol{G}}$ en espacio vectorial $\mathbb{R}^2$, poniendo:

$$s(X) = P + p_{\boldsymbol{F}}(\overrightarrow{PX}) - p_{\boldsymbol{G}}(\overrightarrow{PX}) = s(P) + \tilde{s}(\overrightarrow{PX}),$$

siendo $\tilde{s} = p_{\boldsymbol{F}} - p_{\boldsymbol{G}}$, aplicación lineal; en consecuencia, s es una aplicación afín.

Lo mismo ocurre en el Ejemplo 4.29, donde $s(X)$ es el simétrico del punto X, respecto al plano $\mathcal{F}$ en la dirección de la recta $\boldsymbol{G}$; o en el Ejemplo 4.30, en el que $s(X)$ es el simétrico de X respecto a la recta $\mathcal{G}$ en la dirección del plano $\boldsymbol{F}$.

Estas consideraciones nos dan pie para dar la siguiente definición de simetría en un espacio afín arbitrario.

4.34 Definición.- *Sea $\mathcal{A}$ un espacio afín asociado a un espacio vectorial $\boldsymbol{E}$; dadas dos variedades lineales $\mathcal{F}$ y $\mathcal{G}$ con subespacios directores $\boldsymbol{F}$ y $\boldsymbol{G}$, respectivamente, tal que $\boldsymbol{F} \oplus \boldsymbol{G} = \boldsymbol{E}$ ($\boldsymbol{F}$ y $\boldsymbol{G}$ son suplementarios), y dado un punto $X \in \mathcal{A}$, se llama **simétrico** de X, respecto a $\mathcal{F}$ en la dirección de $\boldsymbol{G}$, al punto $s_{\mathcal{F}}$, tal que*

$$s_{\mathcal{F}}(X) = X + 2\overrightarrow{Xp_{\mathcal{F}}(X)},$$

siendo $p_{\mathcal{F}}$ la proyección de X sobre $\mathcal{F}$ en la dirección de $\boldsymbol{G}$.
*Se dice que $s_{\mathcal{F}}$ es la **simetría** respecto a $\mathcal{F}$ en la dirección del subespacio vectorial $\boldsymbol{G}$.*

Para una simetría s respecto a una variedad lineal $\mathcal{F}$ en la dirección del subespacio vectorial $\boldsymbol{G}$, se tienen unas primeras propiedades, que se deducen directamente de la definición y de las propiedades de las proyecciones:

1) $\mathcal{F}$ es el conjunto de punto fijos de s.

2) $\forall X \in \mathcal{A} \Rightarrow p_{\mathcal{F}}(s(X)) = p_{\mathcal{F}}(X)$.

3) $s^2 = 1_{\mathcal{A}}$.

De estas propiedades surge que "s es una aplicación afín de aplicación lineal asociada $\tilde{s} = p_F - p_G$".

En efecto, si $P \in \mathcal{F}$ se cumple que $s(P) = P$ y

$$s(X) = X + 2\overrightarrow{Xp_{\mathcal{F}}(X)} = P + \overrightarrow{PX} + \overrightarrow{Xp_{\mathcal{F}}(X)} + \overrightarrow{Xp_{\mathcal{F}}(X)} =$$
$$= s(P) + \overrightarrow{Pp_{\mathcal{F}}(X)} + \overrightarrow{Xp_{\mathcal{F}}(X)} = s(P) + p_F(\overrightarrow{PX}) + p_G(\overrightarrow{Xp_{\mathcal{F}}(X)}) =$$
$$= s(P) + p_F(\overrightarrow{PX}) + p_G\big(\overrightarrow{XP} + \overrightarrow{Pp_{\mathcal{F}}(X)}\big) = s(P) + p_F(\overrightarrow{PX}) - p_G(\overrightarrow{PX}) \Rightarrow$$
$$s(X) = s(P) + (p_F - p_G)(\overrightarrow{PX}).$$

La propiedad 3) caracteriza las aplicaciones afines que son proyecciones:

4.34' Definición.- *Una aplicación afín $f \colon \mathcal{A} \to \mathcal{A}$, que verifica $f^2 = 1_{\mathcal{A}}$, se llama **simetría**.*

Para establecer la equivalencia de estas dos definiciones de simetría, solo nos basta, a partir de esta última, encontrar una subvariedad $\mathcal{F}$ y una subespacio vectorial director que verifiquen las condiciones de la Definición 4.6.

Supongamos que $f \colon (\mathcal{A}, E) \to (\mathcal{A}, E)$ es una aplicación afín tal que $f^2 = 1_{\mathcal{A}}$, y sea $\tilde{f}$ su aplicación lineal asociada (que necesariamente verifica $\tilde{f}^2 = 1_E$). Consideremos el conjunto formado por los puntos medios de X y $f(X)$, para todo $X \in \mathcal{A}$, es decir,

$$\mathcal{F} = \{X + \frac{1}{2}\overrightarrow{Xf(X)} \big/ X \in \mathcal{A}\}$$

El conjunto $\mathcal{F}$ será una variedad lineal si $F = \{\overrightarrow{PQ}/P, Q \in \mathcal{F}\}$ es un subespacio vectorial de E. Sean

$$R = P + \frac{1}{2}\overrightarrow{Pf(P)}, \qquad S = Q + \frac{1}{2}\overrightarrow{Qf(Q)},$$
$$\overrightarrow{RS} = \overrightarrow{PQ} + \frac{1}{2}\overrightarrow{Qf(Q)} - \frac{1}{2}\overrightarrow{Pf(P)} =$$
$$= \overrightarrow{PQ} + \frac{1}{2}\overrightarrow{Qf(Q)} - \frac{1}{2}\Big(\overrightarrow{PQ} + \overrightarrow{Pf(P)}\Big) =$$
$$= \frac{1}{2}\Big(\overrightarrow{PQ} + \overrightarrow{f(P)f(Q)}\Big) = \frac{1}{2}\Big(\overrightarrow{PQ} + \tilde{f}(\overrightarrow{PQ})\Big).$$

Luego, F es el subespacio vectorial $\{\vec{v} + \tilde{f}(\vec{v})/\vec{v} \in E\}$ y, por tanto, $\mathcal{F}$ es una variedad lineal.

Ahora consideramos el subconjunto $G = \{\overrightarrow{Xf(X)}/X \in \mathcal{A}\} \subset E$, que va a ser el subespacio vectorial que determine la dirección de la simetría.

Sea $P \in \mathcal{F}$ $(f(P) = P)$, entonces, si $X \in \mathcal{A}$, $\overrightarrow{PX} = \vec{v}$ y

$$\overrightarrow{Xf(X)} = \overrightarrow{(P + \vec{v})f(P + \vec{v})} = \overrightarrow{(P + \vec{v})(f(P) + \tilde{f}(\vec{v}))} =$$
$$= \overrightarrow{(P + \vec{v})(P + \tilde{f}(\vec{v}))} = -\vec{v} + \tilde{f}(\vec{v}) = (\tilde{f} - 1_E)(\vec{v}).$$

Así, $\boldsymbol{G}$ es el subespacio vectorial $\{\tilde{f}(\vec{\boldsymbol{v}}) - \vec{\boldsymbol{v}} / \vec{\boldsymbol{v}} \in \boldsymbol{E}\}$.

Los subespacios $\boldsymbol{F}$ y $\boldsymbol{G}$ son suplementarios:

— $\boldsymbol{F} \cap \boldsymbol{G} = \{\vec{\boldsymbol{0}}_E\}$. Sea $\vec{\boldsymbol{u}} \in \boldsymbol{F} \cap \boldsymbol{G}$, existen $\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2 \in \boldsymbol{E}$ tales que $\vec{\boldsymbol{u}} = \tilde{f}(\vec{\boldsymbol{v}}_1) + \vec{\boldsymbol{v}}_1 = \tilde{f}(\vec{\boldsymbol{v}}_2 - \vec{\boldsymbol{v}}_2$. Entonces, aplicando $\tilde{f}$ a los dos miembros de la última igualdad y usando $\tilde{f}^2 = 1_E$, se obtiene las relaciones:

$$\vec{\boldsymbol{v}}_1 + \vec{\boldsymbol{v}}_2 = \tilde{f}(\vec{\boldsymbol{v}}_2) - \tilde{f}(\vec{\boldsymbol{v}}_1) \qquad \vec{\boldsymbol{v}}_2 - \vec{\boldsymbol{v}}_1 = \tilde{f}(\vec{\boldsymbol{v}}_1) + \tilde{f}(\vec{\boldsymbol{v}}_2).$$

De donde que, restando y sumando miembro a miembro, se tiene:

$$\vec{\boldsymbol{v}}_1 = -\tilde{f}(\vec{\boldsymbol{v}}_1) \qquad \vec{\boldsymbol{v}}_2 = \tilde{f}(\vec{\boldsymbol{v}}_2).$$

Por tanto, $\vec{\boldsymbol{u}} = \vec{o}_E$.

— $\boldsymbol{E} = \boldsymbol{F} + \boldsymbol{G}$. Si $\vec{\boldsymbol{u}} \in \boldsymbol{E}$, $\tilde{f}(\vec{\boldsymbol{u}}) + \vec{\boldsymbol{u}} \in \boldsymbol{F}$ y $\tilde{f}(\vec{\boldsymbol{u}}) - \vec{\boldsymbol{u}} \in \boldsymbol{G}$; con lo que:

$$\vec{\boldsymbol{u}} = \frac{1}{2}(\tilde{f}(\vec{\boldsymbol{u}}) + \vec{\boldsymbol{u}}) + \frac{1}{2}(\vec{\boldsymbol{u}} - \tilde{f}(\vec{\boldsymbol{u}})) \in \boldsymbol{F} + \boldsymbol{G}.$$

Veamos finalmente que la relación entre f y $p_{\mathcal{F}}$ (proyección sobre $\mathcal{F}$ paralelamente a $\boldsymbol{G}$) es la que figura en la Definición 4.6.

Sean $P \in \mathcal{F}$ $(f(P) = P)$ y $X \in \mathcal{A}$,

$$p_{\mathcal{F}}(X) = P + p_F(\overrightarrow{PX}) = P + p_F\Big(\frac{1}{2}(\tilde{f}(\overrightarrow{PX}) + \overrightarrow{PX}) + \frac{1}{2}(\overrightarrow{PX} - \tilde{f}(\overrightarrow{PX}))\Big) =$$

$$= P + \frac{1}{2}(\tilde{f}(\overrightarrow{PX}) + \overrightarrow{PX}) = P + \frac{1}{2}\overrightarrow{f(P)f(X)} + \frac{1}{2}\overrightarrow{PX} = P + \frac{1}{2}\overrightarrow{Pf(X)} + \frac{1}{2}\overrightarrow{PX} \Rightarrow$$

$$\overrightarrow{Xp_{\mathcal{F}}(X)} = \overrightarrow{XP} + \frac{1}{2}\overrightarrow{Pf(X)} - \frac{1}{2}\overrightarrow{XP} \quad \Rightarrow$$

$$2\overrightarrow{Xp_{\mathcal{F}}(X)} = \overrightarrow{XP} + \overrightarrow{Pf(X)} = \overrightarrow{Xf(X)} \quad \Rightarrow$$

$$f(X) = X + 2\overrightarrow{Xp_{\mathcal{F}}(X)}.$$

Las afinidades cuyo cuadrado es la identidad (involutivas), es decir, las simetrías, tienen una matriz asociada (en una referencia adecuada) sencilla. Sea s una simetría respecto a $\mathcal{F}$ en la dirección del subespacio vectorial $\boldsymbol{G}$.

Si $\dim \mathcal{A} = n$, $\dim \boldsymbol{F} = r$ y $\mathcal{R} = \{O; \vec{\boldsymbol{u}}_1, \ldots, \vec{\boldsymbol{u}}_r, \vec{\boldsymbol{v}}_1, \ldots, \vec{\boldsymbol{v}}_{n-r}\}$ es una referencia cartesiana en $\mathcal{A}$ con $O \in \mathcal{F}$, $\vec{\boldsymbol{u}}_i \in \boldsymbol{F}$ $(i = 1, \ldots, r)$ y $\vec{\boldsymbol{v}}_j \in \boldsymbol{G}$ $(j = 1, \ldots, n-r)$, la matriz asociada a s en la referencia $\mathcal{R}$ es:

$$M(s, \mathcal{R}) = \begin{pmatrix} & \overbrace{\qquad\qquad}^{r} & \overbrace{\qquad\qquad}^{n-r} \\ 1 & 0 \ \ldots \ 0 & 0 \ \ldots \ 0 \\ 0 & 1 \ \ldots \ 0 & 0 \ \ldots \ 0 \\ \vdots & \vdots \ \ddots \ \vdots & \vdots \ \ddots \ \vdots \\ 0 & 0 \ \ldots \ 1 & 0 \ \ldots \ 0 \\ 0 & 0 \ \ldots \ 0 & -1 \ \ldots \ 0 \\ \vdots & \vdots \ \ddots \ \vdots & \vdots \ \ddots \ \vdots \\ 0 & 0 \ \ldots \ 0 & 0 \ \ldots \ -1 \end{pmatrix}$$

4.35 Ejemplo.- *La aplicación $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ dada por $f(x,y,z) = (1-x, y, -z)$ es una simetría.*

Se trata de una aplicación afín: Si $X = (x, y, z)$ y $O(0, 0, 0)$, $f(X) = f(O) + \tilde{f}(\overrightarrow{OX})$, con $\tilde{f}(x, y, z) = (-x, y, -z)$. Además, $f^2 = 1_{\mathbb{R}^3}$.

Para determinar la variedad lineal $\mathcal{F}$ sobre la cual se realiza la simetría, debemos determinar su subespacio director:

$$\boldsymbol{F} = \{\tilde{f}(x, y, z) + (x, y, z) / (x, y, z) \in \mathbb{R}^3\} = \{(0, y, 0) / y \in \mathbb{R}\}.$$

Y un punto fijo de f, por ejemplo $P = (1/2, 0, 0)$. O bien, simplemente determinando sus punto fijos:

$$\mathcal{F} = \{x, y, z) \in \mathbb{R}^3 / (1 - x, y, z) = (x, y, z)\} = \{(1/2, y, 0) / y \in \mathbb{R}\}.$$

La dirección según la cual se hace la simetría es la del subespacio vectorial:

$$\boldsymbol{G} = \{\tilde{f}(x, y, z) - (x, y, z) / (x, y, z) \in \mathbb{R}^3\} = \{(-2x, 0, -2z) / x, y \in \mathbb{R}\},$$

es decir, el plano coordenado XOZ.

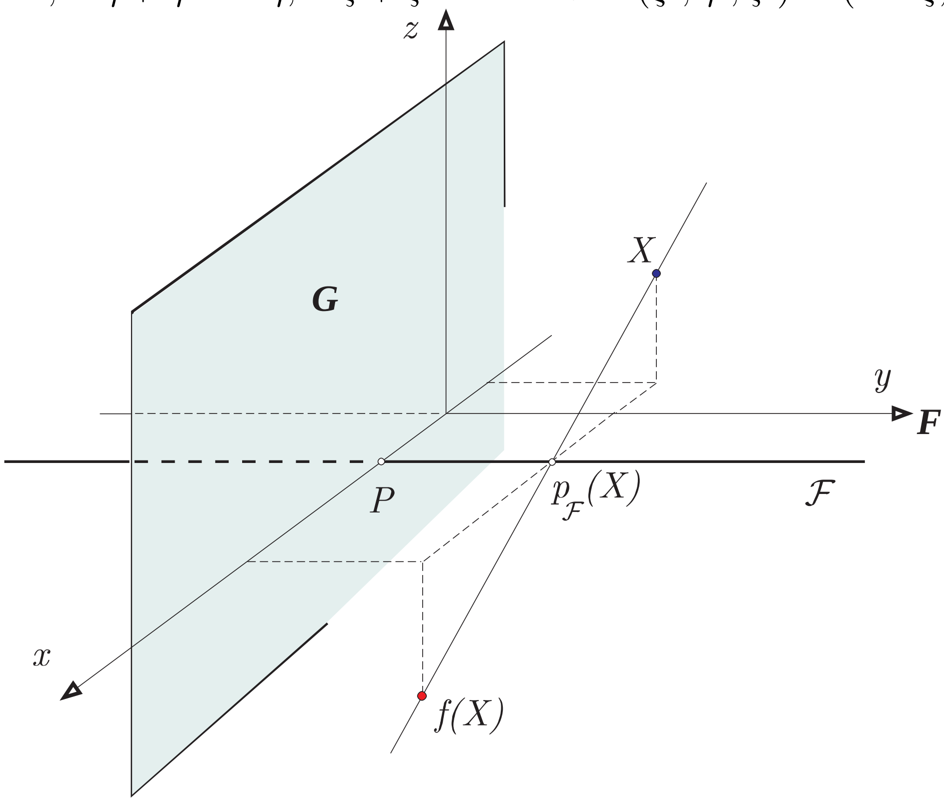
Conociendo la recta $\mathcal{F}$, eje de simetría, y la dirección dada por el plano XOY, podemos determinar las ecuaciones de la simetría en cuestión. Lo cual hacemos por dos vías:

1) A partir de la definición de simetría dada en 4.6. Sean $X = (x, y, z) \in \mathbb{R}^2$ y $P = (1/2, 0, 0) \in \mathcal{F}$, entonces la proyección de X sobre $\mathcal{F}$ en la dirección de $\boldsymbol{G}$ es:

$$p_{\mathcal{F}}(X) = P + p_{\boldsymbol{F}}(\overrightarrow{PX}) = (1/2, 0, 0) + p_{\boldsymbol{F}}(x - 1/2, y, z) =$$
$$= (1/2, 0, 0) + (0, y, 0) = (1/2, y, 0).$$
$$f(X) = X + 2\overrightarrow{Xp_{\mathcal{F}}(X)} = (x, y, z) + 2\big((1/2, y, 0) - (x, y, z)\big) = (1 - x, y, -z).$$

2) Usando los conocimientos sobre rectas y planos en el espacio ordinario. El plano paralelo a XOY por el punto (ξ, η, ζ) tiene por ecuación $y = \eta$; éste corta a la recta $\mathcal{F}$, $x = 1/2, z = 0$ en el punto $(1/2, \eta, 0)$. El simétrico (ξ', η', ζ') de (ξ, η, ζ) respecto a $(1/2, \eta, 0)$ se obtiene de:

$$\xi + \xi' = 1, \quad \eta + \eta' = 2\eta, \quad \zeta + \zeta' = 0 \quad \Rightarrow \quad (\xi', \eta', \zeta') = (1 - \xi, \eta, -\zeta).$$



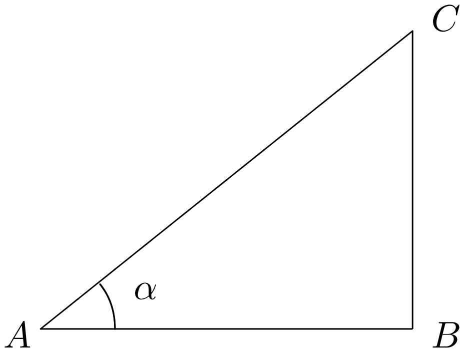
TEMA 5
Espacios vectoriales reales euclídeos

En el espacio tridimensional ordinario $\mathcal{E}$, además de la estructura afín (pág. 18), aparecen las nociones métricas como longitud de segmentos, distancia entre puntos, perpendicularidad, áreas, volúmenes, etc. Para tratar estas nociones se introduce una nueva estructura que consiste en dar una métrica o producto escalar que permite tratar estos conceptos en el marco adecuado.

Para llegar a una definición formal de producto escalar en un espacio vectorial abstracto ahondamos, como hemos hecho antes de introducir las nociones de espacio vectorial (pág. 1) o de espacio afín (pág. 18), en el origen y motivación de la misma, haciendo uso de conocimientos adquiridos en geometría elemental en el espacio ordinario $\mathcal{E}$.

5.1 Caso de los vectores libres del espacio ordinario

Recordar como se definen las funciones trigonométricas seno, coseno y tangente de un ángulo:



$$\operatorname{sen}\alpha = \frac{BC}{AC} \qquad \cos\alpha = \frac{AB}{AC} \qquad \operatorname{tag}\alpha = \frac{BC}{AB}$$

Cada una de estas igualdades relacionan el ángulo α con los lados del triángulo rectángulo $\widehat{ABC}$, de ángulo recto en B.

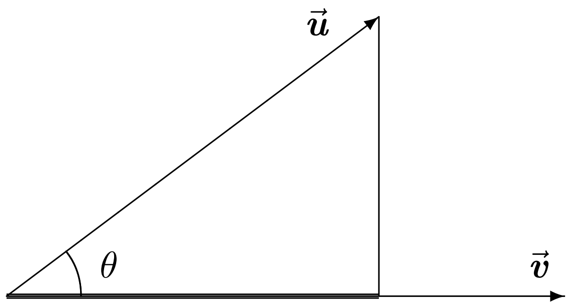
Un vector $\overrightarrow{PQ}$ en el espacio ordinario queda determinado por dos puntos; su origen P y su extremo Q; la distancia entre ellos se llama módulo del vector y se designa por $\|\overrightarrow{PQ}\|$.

Para poder relacionar cómodamente las longitudes de los vectores y los ángulos entre ellos, con sus componentes se da la siguiente definición:

$$\vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}} = \|\vec{\boldsymbol{u}}\| \|\vec{\boldsymbol{v}}\| \ \cos(\widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}}).$$

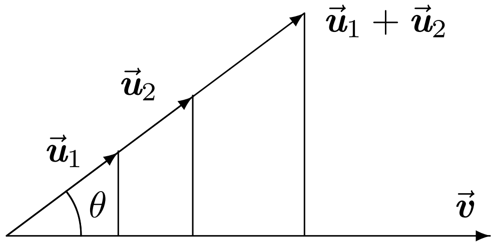
Como $\|\vec{\boldsymbol{u}}\|$, $\|\vec{\boldsymbol{v}}\|$ y $\cos(\widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}})$ son numeros, el resultado es un número, por lo que se da a esta operación con vectores el nombre de "producto escalar".

La proyección de un vector $\vec{\boldsymbol{u}}$ sobre otro $\vec{\boldsymbol{v}}$ se expresa en términos al producto escalar por:



$$\mathrm{p}_{\vec{\boldsymbol{v}}} \vec{\boldsymbol{u}} = \|\vec{\boldsymbol{u}}\| \cos\theta = \frac{\vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}}}{\|\vec{\boldsymbol{v}}\|}$$

Ayudándonos de la noción de proyección de un vector sobre otro se pueden establecer unas propiedades del producto escalar:



$$\begin{array}{ll} 1) & \vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}} = \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{u}} \\ 2) & (\vec{\boldsymbol{u}}_1 + \vec{\boldsymbol{u}}_2) \cdot \vec{\boldsymbol{v}} = \vec{\boldsymbol{u}}_1 \cdot \vec{\boldsymbol{v}} + \vec{\boldsymbol{u}}_2 \cdot \vec{\boldsymbol{v}} \\ 3) & (\lambda \vec{\boldsymbol{u}}) \cdot \vec{\boldsymbol{v}} = \lambda(\vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}}) \\ 4) & \vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{u}} \geq 0 \\ 5) & \vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{u}} = 0 \Leftrightarrow \vec{\boldsymbol{u}} = \vec{o} \end{array} \tag{5-1}$$

Relaciones que se justifican por:
La 1) $\cos(\widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}}) = \cos(\widehat{\vec{\boldsymbol{v}}, \vec{\boldsymbol{u}}})$.
La 2) $\mathrm{p}_{\vec{\boldsymbol{v}}}(\vec{\boldsymbol{u}}_1 + \vec{\boldsymbol{u}}_2) = \mathrm{p}_{\vec{\boldsymbol{v}}} \vec{\boldsymbol{u}}_1 + \mathrm{p}_{\vec{\boldsymbol{v}}} \vec{\boldsymbol{u}}_2$.
La 3) $\|\lambda \vec{\boldsymbol{u}}\| \|\vec{\boldsymbol{v}}\| \ \cos(\widehat{\lambda \vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}}) = \lambda \|\vec{\boldsymbol{u}}\| \|\vec{\boldsymbol{v}}\| \ \cos(\widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}})$.
La 4) $\cos(\widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{u}}}) = 1$.
La 5) $\|\vec{\boldsymbol{u}}\| = 0 \Leftrightarrow \vec{\boldsymbol{u}} = \vec{o}$.

Si $\boldsymbol{E}$ es un espacio vectorial real, tomaremos como punto de partida estas propiedades como axiomas que deben verificar una aplicación $\varphi \colon \boldsymbol{E} \times \boldsymbol{E} \to I\!R$ para que sea un producto escalar. Las propiedades 2) y 3) expresan linealidad en los dos argumentos; por este requisito empezamos para introducir el producto escalar sobre un espacio vectorial, estructura adicional que permitirá formalizar conceptos tales como norma de un vector, ángulo entre vectores, base ortonormal, producto vectorial, etc.

5.2 Producto escalar

El concepto básico para definir una estructura euclídea en un espacio vectorial real [(1)] es el llamado producto escalar de vectores, que es una forma bilineal simétrica con algunas propiedades extras.

5.1 Definición.- *Se llama **forma bilineal** sobre un espacio vectorial real $\boldsymbol{E}$ a una aplicación*

$$\varphi\colon \boldsymbol{E}\times\boldsymbol{E}\to I\!R, \qquad (\vec{\boldsymbol{u}},\vec{\boldsymbol{v}})\mapsto\varphi(\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}),$$

que es lineal en los dos argumentos, es decir:

$$1)\ \varphi(\lambda\vec{\boldsymbol{u}}_1+\mu\vec{\boldsymbol{u}}_2,\vec{\boldsymbol{v}})=\lambda\,\varphi(\vec{\boldsymbol{u}}_1,\vec{\boldsymbol{v}})+\mu\,\varphi(\vec{\boldsymbol{u}}_2,\vec{\boldsymbol{v}}).$$

$$2)\ \varphi(\vec{\boldsymbol{u}},\lambda\vec{\boldsymbol{v}}_1+\mu\vec{\boldsymbol{v}}_2)=\lambda\,\varphi(\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}_1)+\mu\,\varphi(\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}_2).$$

para $\forall\vec{\boldsymbol{u}},\vec{\boldsymbol{u}}_1,\vec{\boldsymbol{u}}_2,\vec{\boldsymbol{v}},\vec{\boldsymbol{v}}_1,\vec{\boldsymbol{v}}_2\in\boldsymbol{E}$ y $\forall\lambda,\mu\in I\!R$.

5.2 Ejemplo.- *Sea $\boldsymbol{E}$ un espacio vectorial sobre un cuerpo K y sean $f\colon\boldsymbol{E}\to K$ y $h\colon\boldsymbol{E}\to K$ aplicaciones lineales. Entonces $\varphi\colon\boldsymbol{E}\times\boldsymbol{E}\to K$ definida por $\varphi(\vec{\boldsymbol{u}},\vec{\boldsymbol{v}})=f(\vec{\boldsymbol{u}})h(\vec{\boldsymbol{v}})$ es una forma bilineal.*

Matriz asociada a una forma bilineal

Si $\mathcal{B}=\{\vec{\boldsymbol{e}}_1,\dots,\vec{\boldsymbol{e}}_n\}$ es una base de $\boldsymbol{E}$, la matriz cuyos coeficientes son los números reales $a_{ij}=\varphi(\vec{\boldsymbol{e}}_i,\vec{\boldsymbol{e}}_j)$, $i,j=1,\dots,n$, se denomina ***matriz de la forma bilineal***.

Así, una forma bilineal $\varphi\colon\boldsymbol{E}\times\boldsymbol{E}\to I\!R$ se puede expresar de la siguiente manera:

$$\varphi(\vec{\boldsymbol{u}},\vec{\boldsymbol{v}})=\sum_{i,j=1}^{n}a_{ij}x^iy^j=\begin{pmatrix}x^1 & x^2 & \cdots & x^n\end{pmatrix}\begin{pmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nn}\end{pmatrix}\begin{pmatrix}y^1\\ y^2\\ \vdots\\ y^n\end{pmatrix}.$$

donde $(x^1,\dots,x^n)$ y $(y^1,\dots,y^n)$ son las componentes de los vectores $\vec{\boldsymbol{u}}$ e $\vec{\boldsymbol{v}}$ respecto de la base $\mathcal{B}$. Las coeficientes de la matriz (a_{ij}) son números reales $\varphi(\vec{\boldsymbol{e}}_i,\vec{\boldsymbol{e}}_j)$, $i,j=1,\dots,n$.

Con notación abreviada, la expresión matricial de una forma bilineal φ se puede poner como $\varphi(\vec{\boldsymbol{u}},\vec{\boldsymbol{v}})={}^{\mathrm{t}}XMY=\sum_{i,j=1}^{n}a_{ij}x^iy^j$, siendo X e Y las matrices columnas formadas por las componentes de los vectores $\vec{\boldsymbol{u}}$ e $\vec{\boldsymbol{v}}$, y $M=(a_{ij})$ la matriz asociada a φ respecto de la base $\mathcal{B}$.

Si respecto a dos bases de $\boldsymbol{E}$, $\mathcal{B}$ y $\mathcal{B}$', M y M' son las matrices asociadas a una forma bilineal $\varphi\colon\boldsymbol{E}\times\boldsymbol{E}\to I\!R$, X e Y y X' e Y' son las matrices columnas

[(1)] En todo este tema, aunque no se especifique, siempre consideraremos **espacios vectoriales sobre el cuerpo de los reales**.

de las componentes de dos vectores $\vec{\boldsymbol{u}}$ e $\vec{\boldsymbol{v}}$ y si, finalmente, $X = AX'$ e $Y = AY'$ son las fórmulas de cambio de bases de $\mathcal{B}$ a $\mathcal{B}'$:

$$\left.\begin{array}{l}\varphi(\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}) = {}^{\mathrm{t}}XMY = {}^{\mathrm{t}}(AX')M(AY') = {}^{\mathrm{t}}X'({}^{\mathrm{t}}AMA)Y' \\ \varphi(\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}) = {}^{\mathrm{t}}X'M'Y'\end{array}\right\} \Rightarrow M' = {}^{\mathrm{t}}AMA. \tag{5-2}$$

5.3 Ejemplo.- *Sea la forma bilineal $\varphi\colon I\!R^3 \times I\!R^3 \to I\!R$ dada por:*

$$\varphi\big((x^1,x^2,x^2),(y^1,y^2,y^3)\big) = (x^1,x^2,x^2)\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} y^1 \\ y^2 \\ y^3\end{pmatrix}.$$

Respecto a la base $\mathcal{B}=\{(-1,-2,1),(1,0,1),(-1,1,1)\}$ la matriz de φ es diagonal.

Pues, el cambio de bases de la canónica $\mathcal{B}_0 = \{(1,0,0),(0,1,0),(0,0,1)\}$ a $\mathcal{B}$ está dado por la matriz:

$$A = \begin{pmatrix} -1 & 1 & -1 \\ -2 & 0 & 1 \\ 1 & 1 & 1\end{pmatrix}.$$

Por lo que $M' = {}^{\mathrm{t}}AMA$ es:

$$\begin{pmatrix} -1 & -2 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 1\end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} -1 & 1 & -1 \\ -2 & 0 & 1 \\ 1 & 1 & 1\end{pmatrix} = \begin{pmatrix} 18 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0\end{pmatrix}.$$

La determinación de los coeficientes de la matriz M' de φ respecto a la base $\mathcal{B} = \{(-1,-2,1),(10,1),(-1,1,1)\}$ se puede hacer a partir de su definición:

$$\begin{pmatrix} \varphi\big((-1,-2,1),(-1,-2,1)\big) & \varphi\big((-1,-2,1),(1,0,1)\big) & \varphi\big((-1,-2,1),(-1,1,1)\big) \\ \varphi\big((1,0,1),(-1,-2,1)\big) & \varphi\big((1,0,1),(1,0,1)\big) & \varphi\big((1,0,1),(-1,1,1)\big) \\ \varphi\big((-1,1,1),(-1,-2,1)\big) & \varphi\big((-1,1,1),(1,0,1)\big) & \varphi\big((-1,1,1),(-1,1,1)\big)\end{pmatrix}.$$

Producto escalar

Para introducir como axioma la propiedad 1) en (5-1) del producto escalar en el espacio afín ordinario, necesitamos la siguiente definición:

5.4 Definición.- *Una forma bilineal $\varphi\colon \boldsymbol{E}\times\boldsymbol{E} \to I\!R$ se dice que es **simétrica** si $\varphi(\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}) = \varphi(\vec{\boldsymbol{v}},\vec{\boldsymbol{u}})$, para todo par de vectores $\vec{\boldsymbol{u}},\vec{\boldsymbol{v}} \in \boldsymbol{E}$.*

5.5 Nota.- Como los coeficientes de la matriz M asociada a una forma bilineal φ respecto a una base $\{\vec{\boldsymbol{e}}_1,\ldots,\vec{\boldsymbol{e}}_n\}$ son $\varphi(\vec{\boldsymbol{e}}_i,\vec{\boldsymbol{e}}_j)$, $i,j = 1,\ldots,n = \dim \boldsymbol{E}$, se tiene una caracterización para que φ sea simétrica: es que M sea simétrica, es decir, $M = {}^{\mathrm{t}}M$.

La otras propiedades 4) y 5) de (5-1), que se deben exigirse a una forma bilineal para que sea producto escalar son las siguientes:

5.6 Definición.- *Diremos que una forma bilineal $\varphi: E \times E \to I\!R$ es **definida** si:*
$$\forall \vec{v} \in E - \{\vec{o}\},\ \varphi(\vec{v}, \vec{v}) \neq 0.$$
*Y φ es **positiva** si:* $\quad \forall \vec{v} \in E,\ \varphi(\vec{v}, \vec{v}) \geq 0.$

5.7 Nota.- Puesto que φ es bilineal, se tiene $\varphi(\vec{o}, \vec{o}) = 0$, y ya que es definida positiva, ocurre que:
$$\varphi(\vec{v}, \vec{v}) = 0 \Leftrightarrow \vec{v} = \vec{o}.$$

Exponemos a continuación un método para averiguar cuando una forma bilineal es definida positiva, haciendo uso de la matriz (a_{ij}) asociada y denotando por:
$$\Delta_k = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}.$$

5.8 Proposición[Criterio de Sylvester].- *Sea E un espacio vectorial real de dimensión finita y M la matriz de una forma bilineal simétrica $\varphi: E \times E \to I\!R$ respecto de una base fijada. Entonces,*
$$\varphi \text{ es definida positiva} \Leftrightarrow \Delta_k > 0,\ (k = 1, \ldots, n).$$

Demostración.- Puede consultarse en [3, Proposicíon 2.3, pág. 254]. ◪

Recopilando las últimas definiciones ya podemos introducir el siguiente concepto:

5.9 Definición.- *Un **producto escalar** en un espacio vectorial real E es una forma bilineal simétrica y definida positiva sobre E:*
$$g: E \times E \to I\!R \qquad (\vec{u}, \vec{v}) \longmapsto g(\vec{u}, \vec{v}) := \vec{u} \cdot \vec{v}.$$

Es decir, verificándose $\forall \vec{u}, \vec{u}_1, \vec{u}_2, \vec{v}, \vec{v}_1, \vec{v}_2 \in E$ y $\forall \lambda, \mu \in I\!R$:
$$\begin{array}{rcl} (\lambda \vec{u}_1 + \mu \vec{u}_2) \cdot \vec{v} & = & \lambda\, \vec{u}_1 \cdot \vec{v} + \mu\, \vec{u}_2 \cdot \vec{v} \\ \vec{u} \cdot (\lambda \vec{v}_1 + \mu \vec{v}_2) & = & \lambda\, \vec{u} \cdot \vec{v}_1 + \mu\, \vec{u} \cdot \vec{v}_2 \\ \vec{u} \cdot \vec{v} & = & \vec{v} \cdot \vec{u} \\ \vec{v} \cdot \vec{v} & \geq & 0 \\ \vec{v} \cdot \vec{v} = 0 & \Leftrightarrow & \vec{v} = \vec{o} \end{array}$$

5.10 Nota.- El producto escalar es también conocido como ***producto punto*** o ***producto interno***.
Cuando una aplicación bilineal de $\varphi: E \times E \to I\!R$ es un producto escalar cambiamos φ por la letra g y para expresar que un espacio vectorial E tiene un producto escalar pondremos $(E, \cdot)$ o bien (E, g).
La matriz asociada al producto escalar la denotamos por (g_{ij}) y la denominaremos ***matriz métrica***.

5.11 Definición.- *Si E tiene dimensión finita, la matriz asociada al producto escalar respecto de una base $\mathcal{B}$ se denomina* ***matriz métrica*** *del producto escalar respecto de la base $\mathcal{B}$.*

5.12 Definición.- *Un* ***espacio vectorial euclídeo*** *es un espacio vectorial real E dotado de un producto escalar.*

5.13 Ejemplo.- *El espacio vectorial real $\mathbb{R}^n$ es euclídeo, con el producto escalar canónico definido para $(x^1,\dots,x^n),(y^1,\dots,y^n)\in\mathbb{R}^n$ por*

$$(x^1,\dots,x^n)\cdot(y^1,\dots,y^n)=x^1y^1+\cdots+x^ny^n$$

Nos referiremos al espacio vectorial euclídeo real canónico $\mathbb{R}^n$ cuando éste está dotado del producto escalar dado en el Ejemplo 5.13.

5.14 Ejemplo.- *Sea E un espacio vectorial real de dimension 2 y $\{\vec{e}_1,\vec{e}_2\}$ una base de E. Si $a>0$ y $b^2-ac<0$, se define el producto escalar por:* $(x^1\vec{e}_1+y^1\vec{e}_2)\cdot(x^2\vec{e}_1+y^2\vec{e}_2)=ax^1x^2+b(x^1y^2+x^2y^1)+cy^1y^2.$

En efecto, se trata de un producto escalar pues las propiedades simétrica y de bilinealidad se verifican fácilmente. Que es definida positiva se deduce aplicando el criterio de Sylvester (Proposición 5.8) a la matriz $\begin{pmatrix} a & b\\ b & c\end{pmatrix}$ asociada a la base $\{\vec{e}_1,\vec{e}_2\}$. En particular:

5.15 Ejemplo.- *La aplicación $\varphi\colon\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}$ definida por:*

$$\varphi((x^1,y^1),(x^2,y^2))=x^1x^2-x^1y^2-y^1x^2+\alpha y^1y^2.$$

es un producto escalar si $\alpha>1$.

5.16 Ejemplo.- *Si $A,B\in\mathcal{M}_{m\times n}(\mathbb{R})$, $A\cdot B=\operatorname{traza}(A\,{}^tB)$ es un producto escalar en $\mathcal{M}_{m\times n}(\mathbb{R})$.*

$$A\cdot B=\operatorname{traza}\left[\begin{pmatrix} a_1^1 & a_1^2 & \cdots & a_1^n\\ \cdots & \cdots & \cdots & \cdots\\ a_m^1 & a_m^2 & \cdots & a_m^n\end{pmatrix}\begin{pmatrix} b_1^1 & \cdots & b_m^1\\ b_1^2 & \cdots & b_m^2\\ \cdots & \cdots & \cdots\\ b_1^n & \cdots & b_m^n\end{pmatrix}\right]=$$

$$(a_1^1b_1^1+\cdots+a_1^nb_1^n)+\cdots+(a_m^1b_m^1+\cdots+a_m^nb_m^n).$$

5.17 Ejemplo.- *Sean $a,b\in\mathbb{R}$, $a<b$ y $\mathcal{C}[a,b]=\{f\colon[a,b]\to\mathbb{R}/f\ \text{continua}\}$ espacio vectorial real (con las operaciones como en el* Ejemplo 1.4*). Se define un producto escalar en $\mathcal{C}[a,b]$ por:*

$$f\cdot h=\int_a^b f(t)h(t)dt.$$

En el caso en que $a = -\pi$ y $b = \pi$ y $n, m \in I\!N$, se pueden determinar los productos escalares:

$$(\operatorname{sen} nt)\cdot(\operatorname{sen} mt) = \begin{cases} 0 \text{ si } m \neq n \\ \pi \text{ si } m = n \end{cases} \qquad (\cos nt)\cdot(\cos mt) = \begin{cases} 0 \text{ si } m \neq n \\ \pi \text{ si } m = n \end{cases}$$

$$(\operatorname{sen} nt)\cdot(\cos mt) = 0, \qquad 1\cdot 1 = \int_{-\pi/2}^{\pi/2} dt = 2\pi, \qquad t\cdot(t^2+2) = 0.$$

5.3 Norma asociada a un producto escalar

5.18 Definición.- *En un espacio vectorial euclídeo $(E,\cdot)$, se llama **norma** o **longitud de un vector** $\vec{v}$ al número real no negativo*

$$\|\vec{v}\| = +\sqrt{\vec{v}\cdot\vec{v}}$$

5.19 Ejemplo.- *La norma asociada al producto escalar euclídeo canónico de $I\!R^n$ (**norma euclídea**) está dada para un vector $(x_1, x_2, \ldots, x_n) \in I\!R^n$ por:*

$$\|(x_1, x_2, \ldots, x_n)\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

5.20 Ejemplo.- *Utilizando la norma asociada al producto escalar en $I\!R^2$ dado por:*
$\varphi((x_1,y_1),(x_2,y_2)) = (x_1,y_1)\cdot(x_2,y_2) = 4x_1x_2 + 2y_1y_2 - 2(x_1y_2 + x_2y_1)$, *la norma del vector $\vec{v} = (2,1)$ es $\|\vec{v}\|_\varphi = \sqrt{10}$, mientras que la norma euclídea del mismo vector es $\|\vec{v}\| = \sqrt{5}$.*

Algunas propiedades de la norma definida a partir de un producto escalar, se describen a continuación.

5.21 Proposición[Propiedades de la norma].- *Dados dos vectores $\vec{u}$ y $\vec{v}$ de un espacio vectorial euclídeo E y un número real λ, se verifica que:*

1) $\|\vec{v}\| = 0 \Leftrightarrow \vec{v} = \vec{0}$.
2) $\|\lambda\vec{v}\| = |\lambda|\|\vec{v}\|$.
3) $\|\vec{u}+\vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$. *(Desigualdad triangular o de Minkowski)*
4) $|\vec{u}\cdot\vec{v}| \leq \|\vec{u}\|\|\vec{v}\|$. *(Desigualdad de Cauchy-Schwarz)*
5) $2(\|\vec{u}\|^2 + \|\vec{v}\|^2) = \|\vec{u}+\vec{v}\|^2 + \|\vec{u}-\vec{v}\|^2$. *(Ley del paralelogramo)*
6) $4\vec{u}\cdot\vec{v} = \|\vec{u}+\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2$.
7) $2\vec{u}\cdot\vec{v} = \|\vec{u}+\vec{v}\|^2 - \|\vec{u}\|^2 - \|\vec{v}\|^2$. *(Identidad de polarización)*
8) $|\|\vec{u}\| - \|\vec{v}\|| \leq \|\vec{u}-\vec{v}\|$.
9) $\|\vec{u}\| = \|\vec{v}\| \Leftrightarrow (\vec{u}+\vec{v})\cdot(\vec{u}-\vec{v})$. *(Ley del rombo)*

Demostración.- Las propiedades 1) e 2) se deducen inmediatamente de la definición de norma asociada a un producto escalar.

4) Supongamos que $\vec{v} \neq \vec{o}$ (para $\vec{v} = \vec{o}$, la verificación es inmediata). Como el producto escalar es definido positivo, para todo $\lambda \in I\!R$, se tiene que:

$$(\vec{u} + \lambda\vec{v}) \cdot (\vec{u} + \lambda\vec{v}) \geq 0.$$

Al ser el producto escalar bilineal y simétrico, esta desigualdad es equivalente a la siguiente:

$$\|\vec{u}\|^2 + 2(\vec{u} \cdot \vec{v})\lambda + \|\vec{v}\|^2\lambda^2 \geq 0.$$

Considerando el primer miembro de esta desigualdad como un polinomio cuadrático en λ, con el coeficiente de λ^2 distinto de cero, no puede tener dos soluciones reales distintas; por lo que:

$$(\vec{u} \cdot \vec{v})^2 - \|\vec{u}\|^2\|\vec{v}\|^2 \leq 0,$$

que es equivalente a la desigualdad que debemos establecer.

Podemos también proceder para demostrar la desigualdad 4) de la forma siguiente:

$$0 \leq \left(\vec{u} - \frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}\vec{v}\right) \cdot \left(\vec{u} - \frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}\vec{v}\right) = \vec{u} \cdot \vec{u} - 2\frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}\vec{u} \cdot \vec{v} + \frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}\frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}\vec{v} \cdot \vec{v} \Rightarrow$$

$$\Rightarrow 0 \leq \|\vec{u}\|^2 - \frac{(\vec{u} \cdot \vec{v})^2}{\|\vec{v}\|^2} \Longrightarrow (\vec{u} \cdot \vec{v})^2 \leq \|\vec{u}\|^2\|\vec{v}\|^2 \Longrightarrow \vec{u} \cdot \vec{v} \leq \|\vec{u}\|\|\vec{v}\|.$$

3) Surge de la desigualdad siguiente y usando 4):

$$\|\vec{u} + \vec{v}\|^2 = (\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v}) = \vec{u} \cdot \vec{u} + 2\vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{v} = \|\vec{u}\|^2 + 2\vec{u} \cdot \vec{u} + \|\vec{v}\|^2 \leq$$
$$\leq \|\vec{u}\|^2 + 2|\vec{u} \cdot \vec{u}| + \|\vec{v}\|^2 \leq \|\vec{u}\|^2 + 2\|\vec{u}\|\|\vec{u}\| + \|\vec{v}\|^2 = (\|\vec{u}\| + \|\vec{v}\|)^2.$$

5) y 6) Desarrollamos $\|\vec{u} + \vec{v}\|^2$ como se hizo al demostrar 3), y también $\|\vec{u} - \vec{v}\|^2$; sólo basta ahora sumar o restar los resultados obtenidos.

7) Esta relación está implícita en la demostración de la propiedad 3).

8) Poniendo $\vec{u} = \vec{u} - \vec{v} + \vec{v}$, tenemos

$$\|\vec{u}\| = \|(\vec{u} - \vec{v}) + \vec{v}\| \leq \|\vec{u} - \vec{v}\| + \|\vec{v}\|,$$

de donde se deduce que: $\|\vec{u}\| - \|\vec{v}\| \leq \|\vec{u} - \vec{v}\|$.

Análogamente, cambiando $\vec{v}$ por $\vec{u}$, tenemos que:

$$\|\vec{v}\| = \|(\vec{v} - \vec{u}) + \vec{u}\| \leq \|\vec{v} - \vec{u}\| + \|\vec{u}\| \Rightarrow \|\vec{v}\| - \|\vec{u}\| \leq \|\vec{v} - \vec{u}\| = \|\vec{u} - \vec{v}\|.$$

De las dos desigualdades obtenidas, resulta que

$$|\|\vec{u}\| - \|\vec{v}\|| \leq \|\vec{u} - \vec{v}\|.$$

9) Sigue inmediatamente de que $(\vec{u} + \vec{v}) \cdot (\vec{u} - \vec{v}) = \|\vec{u}\|^2 - \|\vec{v}\|^2$.

■

La propiedad 9) de la última proposición lo que nos dice es que si $\vec{u}$ y $\vec{v}$ son dos vectores con la misma norma significa que el paralelogramo que ellos determinan es un rombo; es decir, en el paralelogramo las diagonales $\vec{u} + \vec{v}$ y $\vec{u} - \vec{v}$ son perpendiculares.

5.22 Ejemplo.- *La desigualdad de Cauchy–Schwarz* (Proposición 5.21) *aplicada al espacio vectorial euclídeo de las funciones continuas definidas en un intervalo* $[a,b]$ (Ejemplo 5.17) *se expresa en la forma:*

$$\left| \int_a^b f(t)h(t)dt \right| \leq \left(\int_a^b f^2(t)dt \right)^{1/2} \left(\int_a^b h^2(t)dt \right)^{1/2}.$$

5.23 Proposición.- *Para dos vectores* $\vec{u}$ *y* $\vec{v}$ *de un espacio vectorial euclídeo* E *se verifica la igualdad* $|\vec{u}\cdot\vec{v}| = \|\vec{u}\|\|\vec{v}\|$ *si y sólo si* $\vec{u}$ *y* $\vec{v}$ *son linealmente dependientes.*

En efecto, si $\vec{u}$ y $\vec{v}$ son linealmente dependientes, existe $\lambda \in I\!R$ y $\vec{v} = \lambda\vec{u}$, luego

$$|\vec{u}\cdot\vec{v}| = |\vec{u}\cdot\overrightarrow{\lambda \mathbf{u}}| = |\lambda||\vec{u}\cdot\vec{u}| = |\lambda|\|\vec{u}\|^2 = |\lambda|\|\vec{u}\|\|\vec{u}\| = \|\vec{u}\|\|\vec{v}\|.$$

Si $|\vec{u}\cdot\vec{v}| = \|\vec{u}\|\|\vec{v}\|$ entonces la ecuación cuadrática $\|\vec{u}\|^2 + 2(\vec{u}\cdot\vec{v})\lambda + \|\vec{v}\|^2\lambda^2 = 0$ tiene una raíz doble λ_0 y tenemos que $(\vec{u}+\lambda_0\vec{v})\cdot(\vec{u}+\lambda_0\vec{v}) = 0$.

Si $\lambda_0 = 0$, $\|\vec{u}\|^2 = 0 \Rightarrow \vec{u} = \vec{o}$. En este caso, $\vec{u} = \vec{o}$ y $\vec{v}$ son linealmente dependientes.

Si $\lambda \neq 0$, ya que $(\vec{u}+\lambda_0\vec{v})\cdot(\vec{u}+\lambda_0\vec{v}) = 0$, se tiene que $\vec{u}+\lambda_0\vec{v} = \vec{o}$; entonces, $\vec{u}$ y $\vec{v}$ son linealmente dependientes. ◪

5.24 Nota.- En general, se puede definir una norma en un espacio vectorial real E como una función $\|\ \|\colon E \to I\!R$ que cumpla las condiciones 1), 2) y 3) de la Proposición 5.21. Una norma cualquiera en este sentido puede no estar asociada a un producto escalar. Un ejemplo de esto es la norma en $I\!R^2$, definida por $\|(x^1,x^2)\|_1 = |x^1| + |x^2|$. Esta norma no está asociada a ningún producto escalar en $I\!R^2$, pues falla la ley el paralelogramo (5) de Proposición 5.21), como se muestra tomando $\vec{u} = (1,0)$ y $\vec{v} = (0,1)$

$$2(\|\vec{u}\|_1^2 + \|\vec{v}\|_1^2) = 4 \neq 8 = \|\vec{u}+\vec{v}\|_1^2 + \|\vec{u}-\vec{v}\|_1^2.$$

Un producto escalar puede ser recuperado a partir de la norma asociada a él, utilizando la identidad de polarización 7) de la Proposición 5.21:

$$\vec{u}\cdot\vec{v} = \frac{1}{2}\left(\|\vec{u}+\vec{v}\| - \|\vec{u}\| - \|\vec{v}\|\right).$$

De hecho se tiene que:

5.25 Proposición.- *Una norma está asociada a un producto interno si y sólo si la función que se obtiene mediante la identidad de polarización de la norma, es un producto escalar.* ◪

5.26 Proposición.- *Sean* E *es un espacio vectorial real y* $\|\ \|\colon E \to I\!R$ *es una función que cumple con los axiomas de norma y con la identidad de paralelogramo. Entones, la función definida a través de la identidad de polarización es un producto escalar en* E*.* ◪

5.4 Bases ortogonales

5.27 Definición.- *Se dice que un vector de un espacio vectorial euclídeo es* ***unitario*** *si su norma es 1.*

5.28 Definición.- *Se dice que dos vectores $\vec{u}$ y $\vec{v}$ de un espacio vectorial euclídeo son* ***ortogonales*** *y se denota por $\vec{u} \perp \vec{v}$ si su producto escalar $\vec{u} \cdot \vec{v} = 0$.*

5.29 Definición.- *Dada una familia $\{\vec{v}_k\}_{k\in I}$ de vectores en un espacio vectorial euclídeo, se dice que:*
$\{\vec{v}_k\}_{k\in I}$ es ***ortogonal*** *si $\vec{v}_i \cdot \vec{v}_j = 0,\ \forall i, j \in I$ con $i \neq j$.*
$\{\vec{v}_k\}_{k\in I}$ es ***ortonormal*** *si* $\left\{\begin{array}{l} \vec{u}_i \cdot \vec{v}_j = 0,\ \forall i, j \in I \quad i \neq j \\ \|\vec{v}_i\| = \vec{v}_i \cdot \vec{v}_i = 1,\ \forall i \in I \end{array}\right.$.

5.30 Ejemplo.- *La familia $\{\operatorname{sen} nt\}_{n\geq 1} \cup \{\cos mt\}_{m\geq 0}$, de vectores del espacio vectorial euclídeo $\mathcal{C}[-\pi, \pi]$ del* Ejemplo 5.17, *es ortogonal.*

En lo que sigue tratamos de encontrar una base de vectores, respecto a la cual la expresión del producto escalar sea lo mas simple posible, lo que nos va a facilitar el cálculo de las componentes de un vector arbitrario. Antes damos unos resultados que involucran a vectores ortogonales.

5.31 Proposición.- *Toda familia ortogonal de vectores no nulos es linealmente independiente.*

Demostración.- Dado una familia ortogonal de vectores no nulos $\{\vec{v}_1, ..., \vec{v}_r\}$, si $\sum\limits_{i=1}^{r} \lambda^i \vec{v}_i = \vec{0}$, para cada $k \in \{1, \dots, r\}$ se tiene que:

$$0 = \left(\sum_{i=1}^{r} \lambda^i \vec{v}_i\right) \cdot \vec{v}_k = \sum_{i=1}^{r} \lambda^i (\vec{v}_i \cdot \vec{v}_k) = \lambda^k (\vec{v}_k \cdot \vec{v}_k) \Rightarrow \lambda^k = 0.$$

■

Es posible normalizar cualquier familia ortogonal de vectores $\{\vec{v}_1, \dots, \vec{v}_r\}$, no nulos, de un espacio vectorial euclídeo: basta tomar los vectores ortonormales

$$\left\{\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}, \dots, \vec{u}_r = \frac{\vec{v}_r}{\|\vec{v}_r\|}\right\}.$$

5.32 Proposición[Teorema de Pitágoras].- *Dos vectores $\vec{u}$ y $\vec{v}$ de un espacio vectorial euclídeo son ortogonales si y sólo si*

$$\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2.$$

Demostración.- Surge de que $\vec{u}$ y $\vec{v}$ son ortogonales si y sólo si $\vec{u}\cdot\vec{v}=0$, y de

$$\|\vec{u}+\vec{v}\|^2=(\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v})=\|\vec{u}\|^2+2\vec{u}\cdot\vec{v}+\|\vec{v}\|^2.$$

■

5.33 Definición.- *Se denomina* ***base ortogonal*** *de un espacio vectorial euclídeo a una base formada por vectores ortogonales. Se llama* ***base ortonormal*** *a una base formada por vectores ortonormales.*

Es fácil observar que:

• La matriz asociada a un producto escalar respecto a una base ortogonal es diagonal.

• La matriz asociada a un producto escalar respecto a una base ortonormal es la matriz identidad.

5.34 Ejemplo.- *En $\mathbb{R}^3$ el conjunto de vectores*

$$\mathcal{B}=\{\vec{v}_1=(3,4,0),\vec{v}_2=(-4,3,0),\vec{v}_3=(0,0,2)\}$$

son linealmente independientes ($\vec{v}_1\cdot\vec{v}_2=\vec{v}_2\cdot\vec{v}_3=\vec{v}_3\cdot\vec{v}_1=0$). La matriz asociada a un producto escalar canónico en $\mathbb{R}^3$, respecto a esta base ortogonal es $\begin{pmatrix}25&0&0\\0&25&0\\0&0&4\end{pmatrix}$.

La base ortonormal correspondiente a $\mathcal{B}$ es:

$$\left\{\vec{u}_1=\left(\tfrac{3}{5},\tfrac{4}{5},0\right),\vec{u}_2=\left(-\tfrac{4}{5},\tfrac{3}{5},0\right),\vec{u}_3=(0,0,1)\right\}$$

Método de ortogonalización de Gram–Schmidt

5.35 Proposición.- *Si $\mathsf{A}=\{\vec{v}_1,\dots,\vec{v}_r\}$ son vectores linealmente independientes en el espacio vectorial euclídeo E, entonces existen vectores $\mathsf{A}_o=\{\vec{u}_1,\dots,\vec{u}_r\}$ ortogonales tales que si $\mathsf{A}^i=\{\vec{v}_1,\dots,\vec{v}_i\}$ y $\mathsf{A}_o^i=\{\vec{v}_1,\dots,\vec{v}_i\}$ para $1\le i\le r$ se tiene que $\widetilde{\mathsf{A}^i}=\widetilde{\mathsf{A}_o^i}$.*

Demostración.- El proceso para conseguir este conjunto ortogonal, llamado ***método de ortogonalización de Gram-Schmidt***, es el siguiente:

— Primero se toma $\vec{u}_1=\vec{v}_1$.

— A continuación se elige $\vec{u}_2=\vec{v}_2+\lambda_{21}\vec{u}_1$ tal que se cumpla $\vec{u}_1\cdot\vec{u}_2=0$, es decir,

$$\vec{u}_1\cdot\vec{u}_2=\vec{u}_1\cdot\vec{v}_2+\lambda_{21}\vec{u}_1\cdot\vec{u}_1=0\Rightarrow\lambda_{21}=-\frac{\vec{u}_1\cdot\vec{v}_2}{\|\vec{u}_1\|^2}$$

Con lo que
$$\vec{u}_2=\vec{v}_2-\frac{\vec{u}_1\cdot\vec{v}_2}{\|\vec{u}_1\|^2}\vec{u}_1$$

— Se elige ahora $\vec{u}_3=\vec{v}_3+\lambda_{31}\vec{u}_1+\lambda_{32}\vec{u}_2$ tal que $\vec{u}_1\cdot\vec{u}_3=\vec{u}_2\cdot\vec{u}_3=0$, es decir,

$$\vec{u}_1\cdot\vec{u}_3=\vec{u}_1\cdot\vec{v}_3+\lambda_{31}\vec{u}_1\cdot\vec{u}_1+\lambda_{32}\vec{u}_1\cdot\vec{u}_2=0\Rightarrow\lambda_{31}=-\frac{\vec{u}_1\cdot\vec{v}_3}{\|\vec{u}_1\|^2}$$

$$\vec{\boldsymbol{u}}_2\cdot\vec{\boldsymbol{u}}_3=\vec{\boldsymbol{u}}_2\cdot\vec{\boldsymbol{v}}_3+\lambda_{31}\vec{\boldsymbol{u}}_2\cdot\vec{\boldsymbol{u}}_1+\lambda_{32}\vec{\boldsymbol{u}}_2\cdot\vec{\boldsymbol{u}}_2=0\Rightarrow\lambda_{32}=-\frac{\vec{\boldsymbol{u}}_2\cdot\vec{\boldsymbol{v}}_3}{\|\vec{\boldsymbol{u}}_2\|^2}$$

Con lo que queda: $$\vec{\boldsymbol{u}}_3=\vec{\boldsymbol{v}}_3-\frac{\vec{\boldsymbol{u}}_1\cdot\vec{\boldsymbol{v}}_3}{\|\vec{\boldsymbol{u}}_1\|^2}\vec{\boldsymbol{u}}_1-\frac{\vec{\boldsymbol{u}}_2\cdot\vec{\boldsymbol{v}}_3}{\|\vec{\boldsymbol{u}}_2\|^2}\vec{\boldsymbol{u}}_2$$

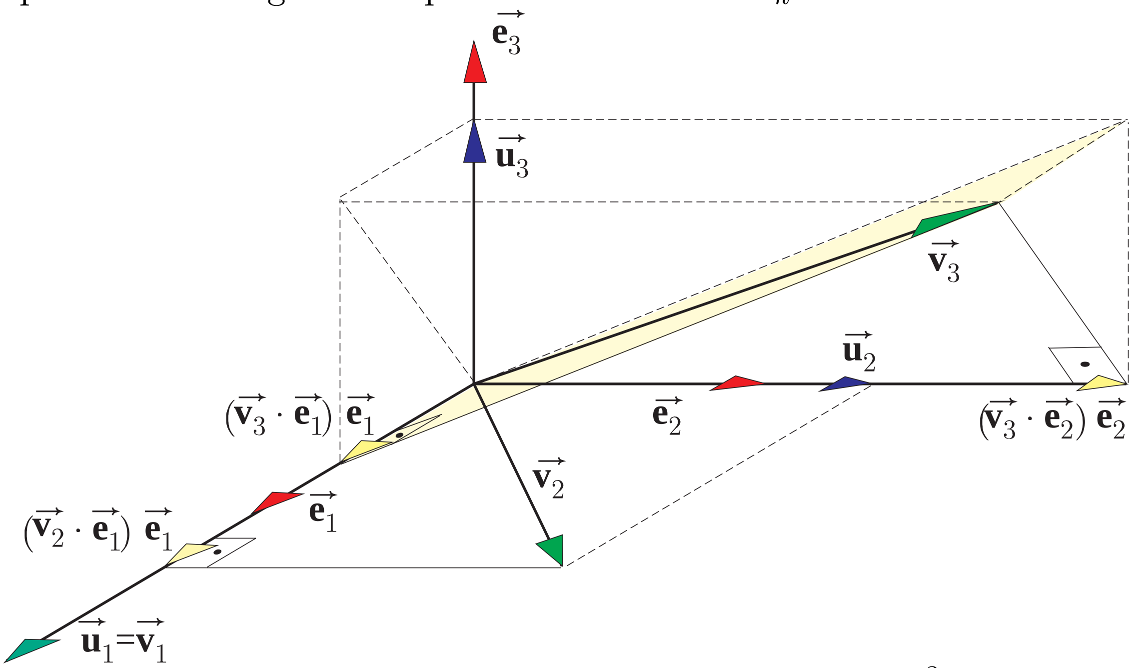
— Continuando este proceso finito, se obtiene que:

$$\vec{\boldsymbol{u}}_k=\vec{\boldsymbol{v}}_k-\sum_{j=1}^{k-1}\frac{\vec{\boldsymbol{u}}_j\cdot\vec{\boldsymbol{v}}_k}{\|\vec{\boldsymbol{u}}_j\|^2}\vec{\boldsymbol{u}}_j\qquad(2\le k\le r)$$

Este método de ortogonalización de Gram–Schmidt nos permite, dada una base en un espacio vectorial euclídeo, obtener una base ortonormal. Una de las características más útiles de las bases ortonormales es que aportan un método muy simple para calcular las coordenadas de un vector a través de cualquier vector de la base. En efecto, supongamos que $\{\vec{\boldsymbol{e}}_1,\dots,\vec{\boldsymbol{e}}_n\}$ es una base ortonormal. Para todo vector $\vec{\boldsymbol{v}}=x^1\vec{\boldsymbol{e}}_1+\cdots+x^n\vec{\boldsymbol{e}}_n$, si calculamos los producto escalares $\vec{\boldsymbol{v}}\cdot\vec{\boldsymbol{e}}_k$ $(1\le k\le n)$ se obtiene

$$\vec{\boldsymbol{v}}\cdot\vec{\boldsymbol{e}}_k=x^1\vec{\boldsymbol{e}}_1\cdot\vec{\boldsymbol{e}}_k+\cdots+x^n\vec{\boldsymbol{e}}_n\cdot\vec{\boldsymbol{e}}_k=x^k\Rightarrow x^k=\vec{\boldsymbol{v}}\cdot\vec{\boldsymbol{e}}_k.$$

Esto significa que $x^k\vec{\boldsymbol{e}}_k=(\vec{\boldsymbol{v}}\cdot\vec{\boldsymbol{e}}_k)\vec{\boldsymbol{e}}_k$ es la proyección ortogonal de $\vec{\boldsymbol{v}}$ sobre el subespacio vectorial generado por el vector básico $\vec{\boldsymbol{e}}_k$.



Construcción de una base ortonormal a partir de la base de $I\!R^3$,
$\{\vec{v}_1=(2,0,0),\vec{v}_2=(3/2,3/2,0),\vec{v}_3=(1/2,3,3/4)\}$.
$\vec{u}_1=\vec{v}_1=(2,0,0)$, $\vec{e}_1=\vec{u}_1/(\vec{u}_1\cdot\vec{u}_1)=(1,0,0)$;
$\vec{u}_2=\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)/(\vec{u}_1\cdot\vec{u}_1)\vec{u}_1=(0,3/2,0)$, $\vec{e}_2=(0,1,0)$;
$\vec{u}_3=\vec{v}_3-(\vec{u}_1\cdot\vec{v}_3)/(\vec{u}_1\cdot\vec{u}_1)\vec{u}_1-(\vec{u}_2\cdot\vec{v}_3)/(\vec{u}_2\cdot\vec{u}_2)\vec{u}_2=(0,0,3/4)$, $\vec{e}_3=(0,0,1)$.

5.36 Ejemplo.- *Sea* $\boldsymbol{F}$ *el subespacio de* $I\!R^4$ *generado por el conjunto de vectores* $\{\vec{\boldsymbol{v}}_1=(1,1,0,0),\vec{\boldsymbol{v}}_2=(1,-1,1,1),\vec{\boldsymbol{v}}_3=(-1,0,2,1)\}$.
Obtener una base ortonormal de $\boldsymbol{F}$.[2]

[2] Con Mathematica, las instrucciones siguiente nos dan la respuesta:
<<LinearAlgebra‘Orthogonalization‘
GramSchmidt[{{1,1,0,0}, {1,-1,1,1}, {-1,0,2,1}}]

Los vectores $\{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2, \vec{\boldsymbol{v}}_3\}$ son linealmente independientes, por lo que forman una base de $\boldsymbol{F}$. Apliquemos el método de Gram-Schmidt:

$$\vec{\boldsymbol{u}}_1 = \vec{\boldsymbol{v}}_1 = (1,1,0,0)$$

$$\vec{\boldsymbol{u}}_2 = \vec{\boldsymbol{v}}_2 - \frac{\vec{\boldsymbol{u}}_1 \cdot \vec{\boldsymbol{v}}_2}{\vec{\boldsymbol{u}}_1 \cdot \vec{\boldsymbol{u}}_1}\vec{\boldsymbol{u}}_1 = \vec{\boldsymbol{v}}_2 - \frac{0}{2}\vec{\boldsymbol{u}}_1 = (1,-1,1,1)$$

$$\vec{\boldsymbol{u}}_3 = \vec{\boldsymbol{v}}_3 - \frac{\vec{\boldsymbol{u}}_1 \cdot \vec{\boldsymbol{v}}_3}{\vec{\boldsymbol{u}}_1 \cdot \vec{\boldsymbol{u}}_1}\vec{\boldsymbol{u}}_1 - \frac{\vec{\boldsymbol{u}}_2 \cdot \vec{\boldsymbol{v}}_3}{\vec{\boldsymbol{u}}_2 \cdot \vec{\boldsymbol{u}}_2}\vec{\boldsymbol{u}}_2 = \vec{\boldsymbol{v}}_3 - \frac{-1}{2}\vec{\boldsymbol{u}}_1 - \frac{2}{4}\vec{\boldsymbol{u}}_2 = \left(-1,1,\frac{3}{2},\frac{1}{2}\right)$$

Dividiendo cada uno de los vectores $\vec{\boldsymbol{u}}_1, \vec{\boldsymbol{u}}_2$ y $\vec{\boldsymbol{u}}_2$ por su norma se obtiene la base ortonormal:

$$\left\{\vec{\boldsymbol{e}}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right), \vec{\boldsymbol{e}}_2 = \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right), \vec{\boldsymbol{e}}_2 = \left(-\frac{\sqrt{2}}{3}, -\frac{\sqrt{2}}{3}, \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{6}\right)\right\}.$$

5.37 Ejemplo.- *Supongamos que en un espacio vectorial euclídeo de dimensión 3 el producto escalar tiene asociada, respecto de una base $\mathcal{B} = \{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2, \vec{\boldsymbol{v}}_3\}$, la matriz métrica siguiente:*

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

Vamos a obtener una base ortonormal respecto a este producto escalar.

Como $\vec{\boldsymbol{v}}_1 \cdot \vec{\boldsymbol{v}}_1 = 0 = a_{11}$, $\vec{\boldsymbol{v}}_3 \cdot \vec{\boldsymbol{v}}_3 = 1 = a_{33}$ y $\vec{\boldsymbol{v}}_1 \cdot \vec{\boldsymbol{v}}_3 = 0 = a_{13}$, estos dos vectores son ortonormales; por lo que sólo nos bastaría encontrar otro vector ortogonal a ellos. Ponemos $\vec{\boldsymbol{u}}_1 = \vec{\boldsymbol{v}}_1$, $\vec{\boldsymbol{u}}_2 = \vec{\boldsymbol{v}}_3$, para determinar $\vec{\boldsymbol{u}}_3$ ortogonal a $\vec{\boldsymbol{u}}_1$ y $\vec{\boldsymbol{u}}_2$ podemos seguir el método de ortogonalización de Gram–Schmidt o bien poner (que en esencia es lo mismo):

$$\vec{\boldsymbol{u}}_3 = \lambda^1\vec{\boldsymbol{v}}_1 + \lambda^2\vec{\boldsymbol{v}}_2 + \lambda^3\vec{\boldsymbol{v}}_3,$$

y encontrar los valores de λ^1, λ^2 y λ^3, tales que $\vec{\boldsymbol{u}}_3 \cdot \vec{\boldsymbol{u}}_1 = 0$ y $\vec{\boldsymbol{u}}_3 \cdot \vec{\boldsymbol{v}}_2 = 0$.

$$\vec{\boldsymbol{u}}_3 \cdot \vec{\boldsymbol{u}}_1 = (\lambda^1\vec{\boldsymbol{v}}_1 + \lambda^2\vec{\boldsymbol{v}}_2 + \lambda^3\vec{\boldsymbol{v}}_3) \cdot \vec{\boldsymbol{v}}_1 = \lambda^1 + \lambda^2 = 0,$$

$$\vec{\boldsymbol{u}}_3 \cdot \vec{\boldsymbol{u}}_2 = (\lambda^1\vec{\boldsymbol{v}}_1 + \lambda^2\vec{\boldsymbol{v}}_2 + \lambda^3\vec{\boldsymbol{v}}_3) \cdot \vec{\boldsymbol{v}}_3 = -\lambda^2 + \lambda^3 = 0.$$

Podemos poner $\lambda^1 = -1, \lambda^2 = 1$ y $\lambda_3 = 1$, y tenemos la base ortogonal.

$$\{\vec{\boldsymbol{u}}_1, \vec{\boldsymbol{u}}_2, \vec{\boldsymbol{u}}_3\} = \{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_3, \vec{\boldsymbol{v}}_1 - \vec{\boldsymbol{v}}_2 + \vec{\boldsymbol{v}}_3\}.$$

La base $\{\vec{\boldsymbol{u}}_1, \vec{\boldsymbol{u}}_2, \vec{\boldsymbol{u}}_3\}$ es ortonormal, pues también $\vec{\boldsymbol{u}}_3$ es unitario:

$$\|\vec{\boldsymbol{u}}_3\| = \vec{\boldsymbol{u}}_3 \cdot \vec{\boldsymbol{u}}_3 = (-1\ 1\ 1)\begin{pmatrix} 1 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = 1.$$

5.38 Nota.- Para demostrar la existencia de vectores ortogonales a partir de un conjunto de ellos y que generen el mismo espacio, hemos utilizado (Proposición 5.35) que se parte de un conjunto finito de ellos, por lo que tal proceso consta de un finito número de pasos. No obstante, se puede proceder por inducción (ver [3, pág. 253]) y, por tanto, el método de Gram–Schmidt es válido también cuando el conjunto de vectores del que se parte es infinito numerable.

5.39 Ejemplo[Polinomios ortogonales].- *En el espacio vectorial de los polinomios en una variable con coeficientes reales, consideramos el producto escalar*

$$f \cdot g = \int_{-1}^{1} f(t)g(t)dt,$$

y aplicamos el método de ortogonalización Gram–Schmidt a los polinomios

$$1, x, x^2, \dots, x^n, \dots$$

para formar una base $\mathcal{B}$ de polinomios ortogonales, relacionados con los Polinomios de Legendre.

Los primeros cuatro elementos de la base ortogonal $\mathcal{B}$, aplicando el método de ortogonalización de Gram–Schmidt, son:

$$P_0(x) = 1, \quad P_1(x) = x - \frac{1 \cdot x}{1 \cdot 1} 1 = x - \frac{\int_{-1}^{1} x dx}{\int_{-1}^{1} dx} = x - \frac{0}{2} = x$$

$$P_2(x) = x^2 - \frac{1 \cdot x^2}{1 \cdot 1} 1 - \frac{x \cdot x^2}{x \cdot x} x = x^2 - \frac{\int_{-1}^{1} x^2 dx}{2} 1 - \frac{\int_{-1}^{1} x^3 dx}{\int_{-1}^{1} x^2 dx} x = x^2 - \frac{1}{3},$$

$$P_3(x) = x^3 - \frac{1 \cdot x^3}{1 \cdot 1} 1 - \frac{x \cdot x^3}{x \cdot x} x - \frac{P_2(x) \cdot x^3}{P_2(x) \cdot P_2(x)} P_2(x) = \frac{1}{5} x(5x^2 - 3).$$

Los cuadrados de las normas de estos cuatro polinomios son:

$$1 \cdot 1 = \int_{-1}^{1} dx = 2, \qquad x \cdot x = \int_{-1}^{1} x^2 dx = \frac{2}{3},$$

$$P_2(x) \cdot P_2(x) = \int_{-1}^{1} \left(x^2 - \frac{1}{3}\right) \cdot \left(x^2 - \frac{1}{3}\right) dx = \frac{8}{45},$$

$$P_3(x) \cdot P_3(x) = \int_{-1}^{1} \left(\frac{1}{5} x(5x^2 - 3)\right)^2 dx = \frac{8}{175}.$$

La expresión de un polinomio general de esta base ortogonal es [3]:

$$P_n(x) = \frac{n!}{(2n)!} \frac{d^n}{dx^n} (x^2 - 1)^n.$$

---

(3) Con Mathematica, las instrucciones siguientes nos da los cinco primeros elementos de una base ortogonal y ortonormal:

In[1]:= <<LinearAlgebra‘Orthogonalization‘

In[2]:= GramSchmidt[{1, x, x^2, x^3, x^4}, InnerProduct → (Integrate[#1*#2,{x,-1,1}]&), Normalized → False]//Simplify

Out[2]:=

$$\left\{1, x, \frac{1}{3}\left(3x^2 - 1\right), \frac{1}{5} x \left(5x^2 - 3\right), \frac{1}{35}\left(35x^4 - 30x^2 + 3\right)\right\}$$

In[3]:= GramSchmidt[{1, x, x^2, x^3, x^4}, InnerProduct → (Integrate[#1*#2,{x,-1,1}]&)]//Simplify

Out[3]:

$$\left\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}} x, \frac{1}{2}\sqrt{\frac{5}{2}}\left(3x^2 - 1\right), \frac{1}{2}\sqrt{\frac{7}{2}} x \left(5x^2 - 3\right), \frac{3\left(35x^4 - 30x^2 + 3\right)}{8\sqrt{2}}\right\}$$

Y del elemento general de la base ortonormal:

$$Q_n(x) = \sqrt{n - (1/2)}P_{n-1}(x).$$

Se trata de los polinomios de Legendre normalizados, los cuales resultan ser precisamente lo que se necesita, por ejemplo, para poder llevar a cabo el análisis del momento angular orbital del electrón en el átomo de hidrógeno, desde el punto de vista de la Mecánica Ondulatoria.
(http://la-mecanica-cuantica.blogspot.com/2009/08/el-espacio-de-hilbert-ii.html 10-6-2011).

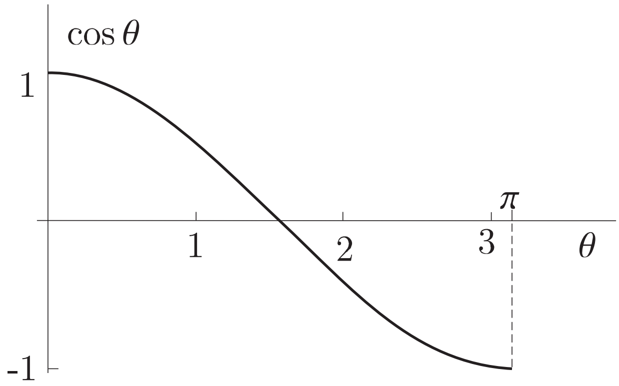
5.5 Ángulo determinado por dos vectores

La desigualdad de Cauchy–Schwarz, es fundamental tanto para probar la desigualdad triangular (3, de Proposición 5.21) de la norma inducida por un producto escalar, como para definir ángulo entre vectores de un espacio vectorial real. La desigualdad de Cauchy–Schwarz equivale a las desigualdades:

$$-1 \leq \frac{\vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}}}{\|\vec{\boldsymbol{u}}\|\|\vec{\boldsymbol{v}}\|} \leq 1.$$

Como la función "cos" es estrictamente decreciente en el intervalo $[0, \pi]$, con valor mínimo $\cos(\pi) = -1$ y valor máximo $\cos(0) = 1$, estas desigualdades no permiten afirmar que si $\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \in \boldsymbol{E} - \{\vec{o}\}$ $((\boldsymbol{E}, \cdot)$ espacio vectorial euclídeo) existe un único número real $\theta \in [0, \pi]$ verificándose que:

$$\cos\theta = \frac{\vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}}}{\|\vec{\boldsymbol{u}}\|\|\vec{\boldsymbol{v}}\|}.$$



5.40 Definición.- *A este valor de θ le denominamos el* ***ángulo*** *determinado por los vectores $\vec{\boldsymbol{u}}$ y $\vec{\boldsymbol{v}}$. Se denotará por $\theta = \widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}}$.*

Obsérvese que el ángulo así definido es un ángulo no orientado, es decir, $\widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}} = \widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}}$.

Damos a continuación algunas propiedades [7, pág. 108] del ángulo determinado por dos vectores, que se deducen de las propiedades del producto escalar que surgen de su Definición 5.9.

5.41 Proposición.- *Si $\vec{\boldsymbol{u}}$ y $\vec{\boldsymbol{v}}$ son vectores no nulos de un espacio vectorial real euclídeo $(\boldsymbol{E}, \cdot)$ y λ es un número real positivo, se tiene:*

1) $\widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}} = \widehat{\vec{\boldsymbol{v}}, \vec{\boldsymbol{u}}}$.

2) $\widehat{\lambda\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}} = \widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}}$.

3) *Si* $\vec{\boldsymbol{u}}$ *y* $\vec{\boldsymbol{v}}$ *son linealmente independientes con* $\|\vec{\boldsymbol{u}}\| = \|\vec{\boldsymbol{v}}\|$, *entonces*
$$\widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{u}}+\vec{\boldsymbol{v}}} = \widehat{\vec{\boldsymbol{u}}+\vec{\boldsymbol{v}}, \vec{\boldsymbol{v}}} \text{ y } \widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{u}}-\vec{\boldsymbol{v}}} = \widehat{\vec{\boldsymbol{u}}-\vec{\boldsymbol{v}}, \vec{\boldsymbol{v}}}.$$

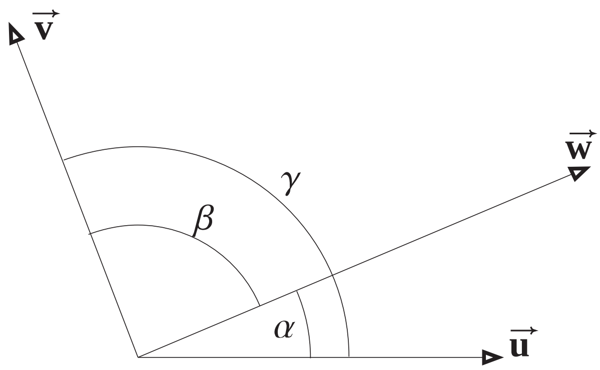
4) $\widehat{-\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}} = \pi - \widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}}$. *(Se dice que los ángulos son* ***suplementarios****)*

5) $\widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}} = 0 \Leftrightarrow \vec{\boldsymbol{v}} = \lambda\vec{\boldsymbol{u}}\ (\lambda > 0)$.

6) $\widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}} = \pi \Leftrightarrow \vec{\boldsymbol{v}} = \lambda\vec{\boldsymbol{u}}\ (\lambda < 0)$.

7) $\widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}} = \pi/2 \Leftrightarrow \vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}} = 0$. *(Los vectores son* ***perpendiculares*** *u* ***ortogonales****)*

8) *Sean* $\vec{\boldsymbol{u}}$ *y* $\vec{\boldsymbol{v}}$ *son linealmente independientes,* λ *y* ν *numeros reales positivos,* $\vec{\boldsymbol{w}} = \lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}$, $\alpha = \widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{w}}}$, $\beta = \widehat{\vec{\boldsymbol{w}}, \vec{\boldsymbol{v}}}$ *y* $\gamma = \widehat{\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}}$. *Entonces*
$$\gamma = \alpha + \beta.$$



5.6 Subespacios ortogonales

En esta sección se estudia la ortogonalidad entre familias de vectores y subespacios vectoriales. Esto será de interés a la hora de tratar la ortogonalidad de subespacios afines (§6.1), así como al trabajar con ciertos tipos de aplicaciones afines (§6.3). Al estar hablando de ortogonalidad, queda implícito que el espacio vectorial en el cual se trabaja es euclídeo. Para empezar estableceremos cuándo dos familias de vectores son ortogonales.

5.42 Definición.- *Dos subconjuntos* A *y* B *de un espacio vectorial euclídeo se dice que son* ***ortogonales*** *(denotado* $\mathsf{A} \perp \mathsf{B}$*) si todo par de vectores* $\vec{\boldsymbol{u}} \in \mathsf{A}$ *y* $\vec{\boldsymbol{v}} \in \mathsf{B}$ *son ortogonales, es decir,* $\vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}} = 0$.

Cuando se tenga un subconjunto de un solo vector $\mathsf{A} = \{\vec{\boldsymbol{u}}\}$ escribiremos $\vec{\boldsymbol{u}} \perp \mathsf{B}$ en lugar de $\{\vec{\boldsymbol{u}}\} \perp \mathsf{B}$.

5.43 Proposición.- *Si* E *es un espacio vectorial euclídeo se verifica que:*

1) *Si* $\mathcal{B}$ *es una base de un subespacio vectorial* $F \dot{\subset} E$ *y* $\vec{\boldsymbol{v}} \in E$, *entonces:*
$$\vec{\boldsymbol{v}} \perp F \Leftrightarrow \vec{\boldsymbol{v}} \perp \mathcal{B}.$$

2) *Si F y G son subespacios vectoriales ortogonales, $F \cap G = \{\vec{o}\}$.*

Demostración.- 1) Para establecer la doble implicación, sólo hay que tener en cuenta que $\mathcal{B} \subset F$ y que todo vector $\vec{v} \in F$ se pone en combinación lineal de los vectores de la base $\mathcal{B}$.

2) Si $\vec{v} \in F \cap G \Rightarrow \vec{v} \in F$ y $\vec{v} \in G \Rightarrow \vec{v} \cdot \vec{v} = 0 \Rightarrow \vec{v} = \vec{o}$.

5.44 Definición.- *Dado un subconjunto* A *de un espacio vectorial euclídeo $(E, \cdot)$, se denomina **subespacio ortogonal** de* A *al conjunto de todos los vectores ortogonales a* A*, esto es:*

$$\mathsf{A}^{\perp} = \{\vec{v} \in E / \vec{v} \cdot \vec{w} = 0, \forall \vec{w} \in \mathsf{A}\}$$

Para que la anterior definición, en la cual se denomina a $\mathsf{A}^{\perp}$ subespacio ortogonal de A, sea coherente, es preciso que el ortogonal de todo subconjunto sea efectivamente subespacio vectorial. Esto es lo que se prueba en el siguiente resultado.

5.45 Proposición.- *Dado un espacio vectorial euclídeo $(E, \cdot)$, el subespacio ortogonal de todo subconjunto de E es un subespacio vectorial.*

Demostración.- Sea $\mathsf{A} \subset E$, como $\vec{o} \in \mathsf{A}^{\perp}$, $\mathsf{A}^{\perp}$ es no vacío.
Además, dados $\lambda, \mu \in \mathbb{R}$ y $\vec{u}, \vec{v} \in \mathsf{A}^{\perp}$, para cualquier $\vec{w} \in \mathsf{A}$ se tiene que:

$$\vec{w} \cdot (\lambda\vec{u} + \mu\vec{v}) = \lambda\vec{w} \cdot \vec{u} + \mu\vec{w} \cdot \vec{v} = 0 \Rightarrow \lambda\vec{u} + \mu\vec{v} \in \mathsf{A}^{\perp}.$$

Pasamos ahora a estudiar casos de espacios ortogonales de subespacios vectoriales.

5.46 Proposición.- *Si F es un subespacio de un espacio vectorial euclídeo $(E, \cdot)$ y $\dim E = n$, se tiene:*

1) *Los subespacios F y $F^{\perp}$ son suplementarios: $E = F \oplus F^{\perp}$.*
2) *Sean $\{\vec{e}_1, \ldots, \vec{e}_r\}$ y $\{\vec{e}_1, \ldots, \vec{e}_r, \vec{e}_{r+1}, \ldots, \vec{e}_n\}$ bases ortogonales de F y de E, respectivamente, entonces $\{\vec{e}_{r+1}, \ldots, \vec{e}_n\}$ es una base de $F^{\perp}$.*
3) $\left(F^{\perp}\right)^{\perp} = F$.
4) $F \subset G \Rightarrow G^{\perp} \subset F^{\perp}$.

Demostración.- Antes de empezar a establecer los apartados de esta proposición, aclarar que la base $\{\vec{e}_1, \ldots \vec{e}_r\}$ del subespacio $F \dot{\subset} E$, se puede completar para formar una base de E [3, pág. 76 Corolario 3.8 (Teorema de Steinitz o de la base incompleta)]. Aplicando luego el método de Gram–Schmidt podemos conseguir una base ortogonal de E de la forma $\{\vec{e}_1, \ldots, \vec{e}_r, \vec{e}_{r+1}, \ldots, \vec{e}_n\}$.

Además, como para $j = r+1, \ldots, n$ los vectores $\vec{e}_j$ son ortogonales a la base $\{\vec{e}_1, \ldots, \vec{e}_r\}$ de F, ocurre (por 1) de la Proposición 5.43) que $\{\vec{e}_{r+1}, \ldots, \vec{e}_n\} \subset F^{\perp}$. Esto nos dice que $\dim F^{\perp} \geq n - r$, pero como $F \cap F^{\perp} = \{\vec{o}\}$ (por 2) de

la Proposición 5.43), $\{\vec{e}_{r+1}, \ldots, \vec{e}_n\}$ es una base de $F^\perp$ y, por tanto, $\dim F^\perp = n - r$.

Por lo expuesto, quedan establecidos los dos primeros apartados, 1) y 2).

3) $F^\perp \oplus \left(F^\perp\right)^\perp = E$ y $F^\perp \oplus F = E \Rightarrow F = \left(F^\perp\right)^\perp$.

5.47 Proposición.- *Sean F y G dos subespacio de un espacio vectorial euclídeo E, entonces*

$$1)\ (F+G)^\perp = F^\perp \cap G^\perp, \qquad 2)\ (F \cap G)^\perp = F^\perp + G^\perp.$$

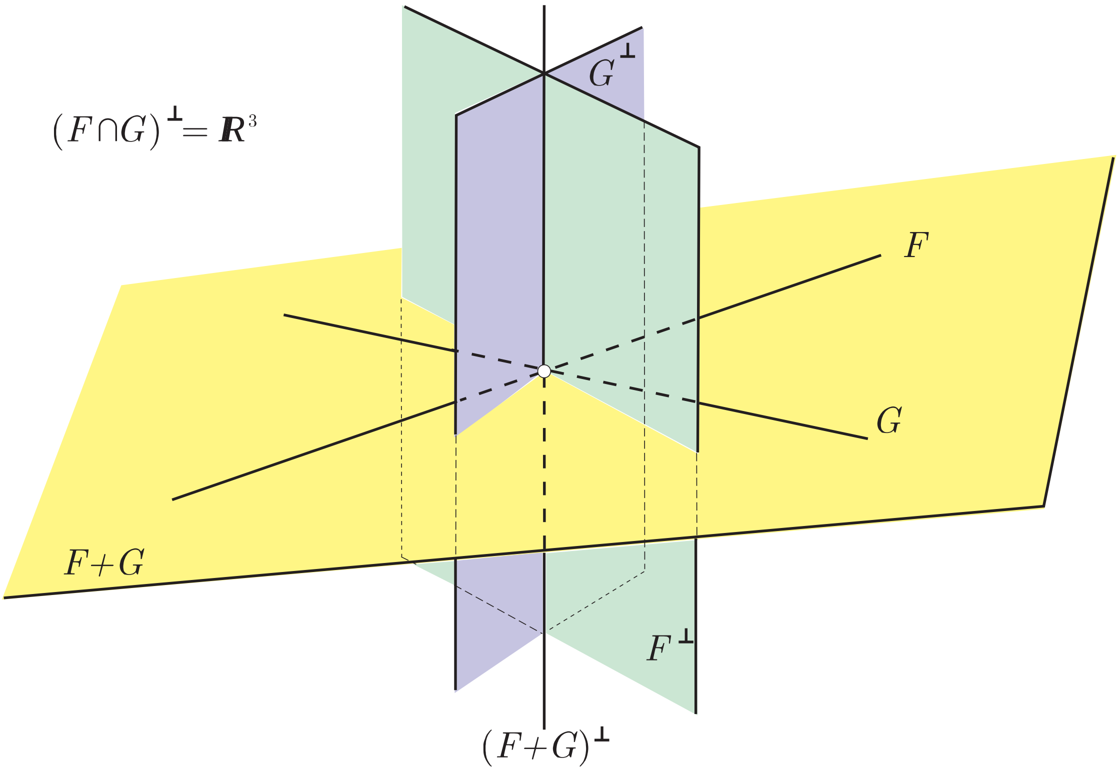
Demostración.- 1) Si un vector $\vec{w}$ es ortogonal a $F+G$, como $F \subset F+G$ y $G \subset F+G$, resulta que $\vec{w}$ es ortogonal a F y a G.

Para lo otra inclusión: un vector de $F+G$ se puede poner de la forma $\vec{u}+\vec{v}$ con $\vec{u} \in F$ y $\vec{v} \in G$.

Así, si $\vec{w} \in F^\perp \cap G^\perp \Rightarrow \vec{w} \cdot (\vec{u}+\vec{v}) = \vec{w} \cdot \vec{u} + \vec{w} \cdot \vec{v} = 0+0 = 0 \Rightarrow \vec{w} \in (F+G)^\perp$.

2) Por la relación 1), se tiene que $(F^\perp + G^\perp)^\perp = (F^\perp)^\perp \cap (G^\perp)^\perp = F \cap G \Rightarrow (F \cap G)^\perp = F^\perp + G^\perp$.

Ejemplo gráfico en $\mathbb{R}^3$:



5.48 Ejemplo.- *Sean $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $\mathcal{F}(A)\dot{\subset}\mathbb{R}^n$ el subespacio generado por los m vectores cuyas componentes son las m filas de A, $\mathcal{C}(A)\dot{\subset}\mathbb{R}^m$ el subespacio generado por los n vectores cuyas componentes son las n columnas de A, es decir,*

$$\mathcal{C}(A) = \{Y \in \mathbb{R}^m / \exists X \in \mathbb{R}^n,\ AX = Y\}$$

y $\mathcal{N}(A)\dot{\subset}\mathbb{R}^n$ es el subespacio de $\mathbb{R}^n$ formado por los vectores que son solución del sistema homogéneo que tiene a A como matriz de coeficientes, es decir,

$$\mathcal{N}(A) = \{X \in \mathbb{R}^n / AX = 0\}.$$

Entonces, respecto al producto escalar euclídeo canónico:

$$\mathcal{N}(A) = (\mathcal{F}(A))^{\perp}, \qquad \mathcal{N}({}^tA) = (\mathcal{C}(A))^{\perp}.$$

5.7 Proyecciones y simetrías ortogonales

Una situación particular de proyectores (Definición 1.39) se plantea en el caso de subespacios vectoriales ortogonales.

Si $F \dot{\subset} E$ un subespacio vectorial del espacio vectorial euclídeo E, cada vector $\vec{u} \in E$ se descompone de forma única (Proposición 5.46) como suma de un vector de $\vec{v} \in F$ y otro de $\vec{w} \in F^{\perp}$, que se llaman, respectivamente, proyección ortogonal de $\vec{u}$ sobre F y sobre $F^{\perp}$, y se denotan:

$$\vec{v} = \mathrm{p}_F\, \vec{u}, \quad \vec{w} = \mathrm{p}_{F^\perp}\, \vec{u}, \qquad (\vec{u} = \mathrm{p}_F\, \vec{u} + \mathrm{p}_{F^\perp}\, \vec{w})$$

Se define el ***ángulo*** entre un vector $\vec{u} \in E$ y un subespacio $F \dot{\subset} E$ como el ángulo que forma con su proyección, es decir:

$$\widehat{\vec{u}, F} = \widehat{\vec{u}, \mathrm{p}_F\, \vec{u}} = \arccos \frac{\vec{u} \cdot \mathrm{p}_F\, \vec{u}}{\|\vec{u}\| \|\mathrm{p}_F\, \vec{u}\|}.$$

Vamos a hallar la proyección de un vector $\vec{u} \in E$ sobre un subespacio F del que se conoce una base.

Cuando $\{\vec{w}_1, \ldots, \vec{w}_r\}$ es una base ortogonal de F:

$$\mathrm{p}_F\, \vec{u} = \sum_{i=1}^{r} (\vec{u} \cdot \vec{w}_i)\vec{w}_i.$$

Si $\{\vec{v}_1, \ldots, \vec{v}_r\}$ es una base arbitraria de F y $\mathrm{p}_F\, \vec{u} = \xi^1 \vec{v}_1 + \cdots + \xi^r \vec{v}_r$, debemos imponer que el vector $\vec{u} - \mathrm{p}_F\, \vec{u}$ sea ortogonal a F, es decir:

$$\left(\vec{u} - \sum_{i=1}^{r} \xi^i \vec{v}_i \right) \cdot \vec{v}_k = 0 \qquad (k = 1, \ldots, r)$$

Resolviendo este sistema de ecuaciones de r ecuaciones y r incógnitas en las variables $\xi^1, \ldots, \xi^r$, encontramos las componentes de $\mathrm{p}_F\, \vec{u}$ respecto a la base de F elegida. Notar que este sistema tiene solución única, ya que la matriz de los coeficientes de este sistema, $(\vec{v}_i \cdot \vec{v}_j)_{i,j=1,\ldots,r}$, es la matriz métrica del producto escalar de E, restringido a F y, por tanto, tiene determinante distinto de cero.

5.49 Ejemplo.- *En $\mathbb{R}^3$ con la estructura euclídea canónica, la proyección (ortogonal) del vector $\vec{u} = (0, 0, 1)$ sobre el subespacio vectorial $F = \{(x, y, z) \in \mathbb{R}^3 / x - y - z = 0\}$ es el vector $\vec{w} = (1/3, -1/3, 2/3)$.*

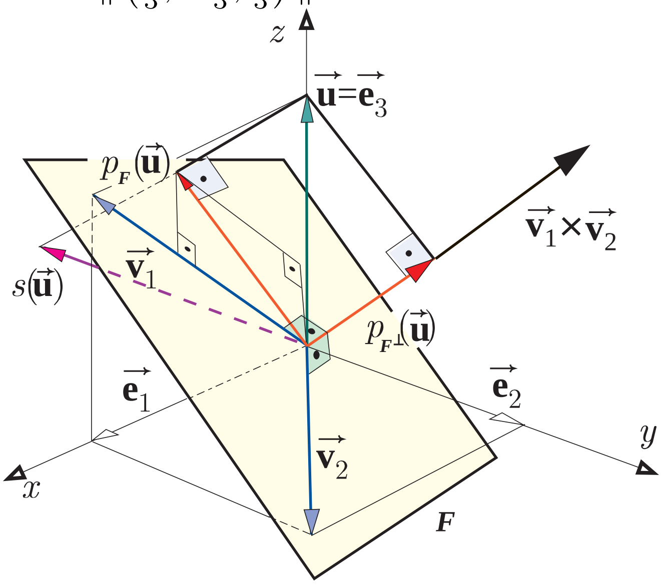
Dos vectores independientes de F son $\vec{v}_1 = (1, 0, 1)$ y $\vec{v}_2 = (1, 1, 0)$. Para hallar la proyección de $\vec{u}$ sobre F, que será de la forma $\lambda\vec{v}_1 + \mu\vec{v}_2$, debemos encontrar λ y μ, para que $\vec{u} - (\lambda\vec{v}_1 + \mu\vec{v}_2)$ sea ortogonal a F.

$$\left.\begin{array}{l} ((0,0,1) - (\lambda(1,0,1) + \mu(1,1,0))) . (1,0,1) = 0 \\ ((0,0,1) - (\lambda(1,0,1) + \mu(1,1,0))) . (1,1,0) = 0 \end{array}\right\} \Rightarrow \left.\begin{array}{l} 2\lambda + \mu = 1 \\ \lambda + 3\mu = 0 \end{array}\right\} \Rightarrow$$

$$\lambda = \frac{2}{3}, \quad \mu = -\frac{1}{3} \Rightarrow \mathrm{p}_F\, \vec{u} = \frac{2}{3}\vec{v}_1 - \frac{1}{3}\vec{v}_2 = \left(\frac{1}{3}, -\frac{1}{3}, \frac{2}{3}\right).$$

$$\mathrm{p}_{F^\perp}\,\vec{\boldsymbol{u}} = \vec{\boldsymbol{u}} - \mathrm{p}_F\,\vec{\boldsymbol{u}} = \left(-\frac{1}{3},\frac{1}{3},\frac{1}{3}\right).$$

$$\widehat{\vec{\boldsymbol{u}},\mathrm{p}_F\,\vec{\boldsymbol{u}}} = \frac{(0,0,1)\cdot\left(\frac{1}{3},-\frac{1}{3},\frac{2}{3}\right)}{\|\left(\frac{1}{3},-\frac{1}{3},\frac{2}{3}\right)\|} = \arccos\sqrt{\frac{2}{3}} \backsimeq 35.2644°.$$



La misma situación se presenta para las simetrías (Definición **??**) en el caso de subespacios ortogonales y suplementarios.

5.50 Definición.- *Sea E un espacio vectorial euclídeo de dimensión finita, para dos subespacios F y G, si F y G forman una suma directa $E = F\oplus G$ y F y G son ortogonales (es decir, $G = F^\perp$), la* `simetría ortogonal` *con respecto a F y paralela a G es la aplicación lineal definida, para todo $\vec{\boldsymbol{v}}\in E$, por*

$$s(\vec{\boldsymbol{v}}) = 2p_F(\vec{\boldsymbol{v}}) - \vec{\boldsymbol{v}}.$$

Cuando F es un hiperplano, decimos que s es la reflexión en F.

En el Ejemplo 5.49:

$$s(\vec{\boldsymbol{u}}) = 2\mathrm{p}_F\,\vec{\boldsymbol{u}} - \vec{\boldsymbol{u}} = \left(\frac{2}{3},-\frac{2}{3},\frac{1}{3}\right).$$

5.8 Producto vectorial

En este párrafo introduciremos el producto vectorial en un espacio vectorial euclídeo. Este concepto nos será de mucha utilidad sobre todo es espacios vectoriales euclídeos tridimensionales, pues permite calcular de una forma rápida y sencilla vectores ortogonales a dos vectores dados. Para este fin es preciso definir previamente la noción de orientación para las bases de un espacio vectorial real.

Orientación de un espacio vectorial euclídeo

5.51 Definición.- *Sean $\mathbf{E}$ un espacio vectorial real de dimensión n, $\mathcal{B} = \{\vec{\boldsymbol{u}}_1, \dots, \vec{\boldsymbol{u}}_n\}$ y $\mathcal{B}' = \{\vec{\boldsymbol{u}}'_1, \dots, \vec{\boldsymbol{u}}'_n\}$ dos bases ordenadas de $\mathbf{E}$ y M la matriz cambio de bases de $\mathcal{B}$ a $\mathcal{B}'$; diremos que las dos bases $\mathcal{B}$ y $\mathcal{B}'$ tienen la **misma orientación** si $|M| := \det(M) > 0$.*
*Diremos que $\mathcal{B}$ y $\mathcal{B}'$ tienen **orientación opuesta** u **contraria** si $|M| < 0$.*

Existen dos orientaciones posibles en $\mathbf{E}$. Dar una ***orientación*** en un espacio vectorial es elegir una de las dos orientaciones posibles. La orientación elegida diremos que es ***positiva*** y la otra orientación, ***negativa***.

5.52 Ejemplo.- *La orientación dada por una base $\{\vec{\boldsymbol{u}}_1, \vec{\boldsymbol{u}}_2, \dots, \vec{\boldsymbol{u}}_n\}$ es opuesta a la dada por las bases $\{\vec{\boldsymbol{u}}_2, \vec{\boldsymbol{u}}_1, \dots, \vec{\boldsymbol{u}}_n\}$ y $\{\vec{\boldsymbol{u}}_1, \dots, \vec{\boldsymbol{u}}_{n-1}, -\vec{\boldsymbol{u}}_n\}$.*

Matriz cambio de bases ortonormales

Vamos a expresar la condición que debe satisfacer la matriz cambio de bases ortonormales.

Sea $\mathcal{B} = \{\vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$ y $\mathcal{B}_0 = \{\vec{\boldsymbol{u}}_1, \dots, \vec{\boldsymbol{u}}_n\}$ dos bases de un espacio vectorial $\mathbf{E}$ de dimensión n. Denotamos por M la matriz cambio de bases de $\mathcal{B}_0$ a $\mathcal{B}$, es decir, $Y = MX$, siendo X e Y las matrices columna de las componentes de un vector respecto a las bases $\mathcal{B}$ y $\mathcal{B}_0$, respectivamente. Si G es la matriz métrica del producto escalar respecto a la base $\mathcal{B}$, la matriz métrica del producto escalar respecto a la base $\mathcal{B}_0$ es (ver 5-2) $G_0 = {}^{\mathrm{t}}MGM$.

Se tiene entonces, al ser la matriz identidad la matriz métrica respecto a una base ortonormal, que:

$$\mathcal{B}_0 \text{ es una base ortonormal} \Longleftrightarrow {}^{\mathrm{t}}MGM = I$$
$$\text{Si } \mathcal{B} \text{ es ortonormal} \Longrightarrow (\mathcal{B}_0 \text{ es ortonormal} \Longleftrightarrow {}^{\mathrm{t}}MM = I \Longleftrightarrow M^{-1} = {}^{\mathrm{t}}M)$$

En consecuencia, el determinante de la matriz cambio de bases ortonormales es ± 1.

5.53 Nota.- Una matriz $A = (a_i^j)_{i,j=1,\dots,n}$, cuadrada con determinante no nulo, cuya inversa es igual a su traspuesta recibe el nombre de ***matriz ortogonal*** o lo que es equivalente el producto por su traspuesta es la matriz identidad, ${}^{\mathrm{t}}AA = I_n$. En términos de sus coeficientes, esta condición se escribe:

$$\sum_{k=1}^{n} a_i^k a_k^j = \delta_i^j = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

Esto quiere decir que las filas (también las columnas) de una matriz ortogonal son las componentes de n vectores que forman una base ortonormal en $\mathbb{R}^n$, respecto al producto escalar canónico. Debido a este hecho debería recibir el nombre de matriz ortonormal; no obstante, una justificación del nombre de matriz ortogonal podría ser que las matrices de una transformación ortogonal son de este tipo (Corolario 5.71).

5.54 Ejemplo.- *Sean el espacio vectorial euclídeo canónico $\mathbb{R}^3$ y la aplicación $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ definida por $(x, y, z) \mapsto (x + 2z, -y, 2x + z)$. Encontrar una base ortonormal respecto a la cual la matriz asociada f sea diagonal.*

Producto vectorial

5.55 Definición.- *Sean $(E, \cdot)$ un espacio vectorial euclídeo real, $\dim \boldsymbol{E} = n$ y $\mathcal{B}_0 = \{\vec{\boldsymbol{e}}_1, \dots, \vec{\boldsymbol{e}}_n\}$ una base ortonormal, se llama **producto vectorial** de $n-1$ vectores $\vec{\boldsymbol{v}}_1, \dots, \vec{\boldsymbol{v}}_{n-1}$, respecto a la orientación determinada por $\mathcal{B}_0$, al vector, denotado por $\vec{\boldsymbol{v}}_1 \times \cdots \times \vec{\boldsymbol{v}}_{n-1}$, tal que para todo $\vec{\boldsymbol{u}} \in \boldsymbol{E}$, verifica*

$$\vec{\boldsymbol{v}}_1 \times \cdots \times \vec{\boldsymbol{v}}_{n-1} \cdot \vec{\boldsymbol{u}} = \begin{vmatrix} u^1 & u^2 & \cdots & u^n \\ v_1^1 & v_1^2 & \cdots & v_1^n \\ \vdots & \vdots & \ddots & \vdots \\ v_{n-1}^1 & v_{n-1}^2 & \cdots & v_{n-1}^n \end{vmatrix}, \tag{5-3}$$

donde $(u^1, u^2, \dots, u^n)$ y $(v_k^1, v_k^2, \dots, v_k^n)$ son las componentes de $\vec{\boldsymbol{u}}$ y $\vec{\boldsymbol{v}}_k$ $(k = 1, \dots, n-1)$ respecto a la base $\mathcal{B}_0$. Podemos denotar el segundo miembro de la fórmula (5-3) por $\det_{\mathcal{B}_0}(\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}_1, \dots, \vec{\boldsymbol{v}}_{n-1})$.

El producto vectorial $\vec{\boldsymbol{v}}_1 \times \cdots \times \vec{\boldsymbol{v}}_{n-1}$ está bien definido, en el sentido de que es el único vector que verifica (5-3) y que es independiente de la base ortonormal que determina la orientación tomada en $\boldsymbol{E}$.

Para lo primero, basta observar que las componentes de $\vec{\boldsymbol{v}}_1 \times \cdots \times \vec{\boldsymbol{v}}_{n-1}$, respecto a la base $\mathcal{B}_0$, sólo dependen de los vectores $\vec{\boldsymbol{v}}_1, \dots, \vec{\boldsymbol{v}}_{n-1}$, es decir, no depende de $\vec{\boldsymbol{u}}$.

La componente i–ésima del producto vectorial $\vec{\boldsymbol{v}}_1 \times \cdots \times \vec{\boldsymbol{v}}_{n-1}$, respecto a la base $\mathcal{B}_0$, es

$$(-1)^{i+1} \begin{vmatrix} v_1^1 & \cdots & v_1^{i-1} & v_1^{i+1} & \cdots & v_1^n \\ v_2^1 & \cdots & v_2^{i-1} & v_2^{i+1} & \cdots & v_2^n \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ v_{n-1}^1 & \cdots & v_{n-1}^{i-1} & v_{n-1}^{i+1} & \cdots & v_1^n \end{vmatrix}$$

Para establecer que el producto vectorial es el mismo cuando se toman dos bases ortonormales con la misma orientación, sea $\mathcal{B}'_0 = \{\vec{\boldsymbol{e}}'_1, \dots, \vec{\boldsymbol{e}}'_n\}$ una base ordenada ortonormal tal que $\det(A) > 0$, siendo $A = (a_i^j)$ la matriz cambio de bases de $\mathcal{B}_0$ a $\mathcal{B}'_0$, la cual es ortogonal: $A^{-1} = {}^{\mathrm{t}}A$ y, por tanto, $\det(A) = 1$.

Si $(u'^1, u'^2, \dots, u'^n)$ y $(v'^1_k, v'^2_k, \dots, v'^n_k)$ son las componentes de $\vec{\boldsymbol{u}}$ y $\vec{\boldsymbol{v}}_k$ $(k = 1, \dots, n-1)$ respecto a la base $\mathcal{B}'_0$, se tiene que $U = AU'$ y $V_k = AV'_k$, siendo U, U', V_k, V'_k matrices columna, formadas por las componentes de los

vectores $\vec{\boldsymbol{u}}$ y $\vec{\boldsymbol{v}}_k$ respecto a las bases correspondientes. La independencia del producto vectorial de las bases $\mathcal{B}_0$ y $\mathcal{B}'_0$ surge, entonces, de que:

$$\begin{pmatrix} u^1 & v_1^1 & \cdots & v_{n-1}^1 \\ u^2 & v_1^2 & \cdots & v_{n-1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ u^n & v_1^n & \cdots & v_{n-1}^n \end{pmatrix} = \begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \cdots & a_n^n \end{pmatrix} \begin{pmatrix} u'^1 & v'^1_1 & \cdots & v'^1_{n-1} \\ u'^2 & v'^2_1 & \cdots & v'^2_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ u'^n & v'^n_1 & \cdots & v'^n_{n-1} \end{pmatrix}.$$

$$\det_{\mathcal{B}_0}(\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}_1, \ldots, \vec{\boldsymbol{v}}_{n-1}) = \det_{\mathcal{B}'_0}(\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}_1, \ldots, \vec{\boldsymbol{v}}_{n-1}).$$

En el caso particular de que $\dim \boldsymbol{E} = 3$ las componentes del producto vectorial de $\vec{\boldsymbol{u}} = u^1\vec{\boldsymbol{e}}_1 + u^2\vec{\boldsymbol{e}}_2 + u^3\vec{\boldsymbol{e}}_3$ y $\vec{\boldsymbol{v}} = v^1\vec{\boldsymbol{e}}_1 + v^2\vec{\boldsymbol{e}}_2 + v^3\vec{\boldsymbol{e}}_3$ son:

$$\vec{\boldsymbol{u}}\times\vec{\boldsymbol{v}} = \begin{vmatrix} u^2 & u^3 \\ v^2 & v^3 \end{vmatrix} \vec{\boldsymbol{e}}_1 - \begin{vmatrix} u^1 & u^3 \\ v^1 & v^3 \end{vmatrix} \vec{\boldsymbol{e}}_2 + \begin{vmatrix} u^1 & u^2 \\ v^1 & v^2 \end{vmatrix} \vec{\boldsymbol{e}}_3 =$$

$$= (u^2v^3 - u^3v^2)\vec{\boldsymbol{e}}_1 + (u^3v^1 - u^1v^3)\vec{\boldsymbol{e}}_2 + (u^1v^2 - u^2v^1)\vec{\boldsymbol{e}}_3. \qquad (5\text{-}4)$$

Tradicionalmente, sobre todo en la literatura de Física, como recurso nemotécnico se suele usar la notación para producto vectorial, cuando $n = 3$, siguiente:

$$\vec{\boldsymbol{u}}\times\vec{\boldsymbol{v}} = \begin{vmatrix} \vec{\boldsymbol{e}}_1 & \vec{\boldsymbol{e}}_2 & \vec{\boldsymbol{e}}_3 \\ u^1 & u^2 & u^3 \\ v^1 & v^2 & v^3 \end{vmatrix},$$

entendiéndose que se trata de desarrollar formalmente este determinante por la primera fila.

Vamos a mostrar una serie de propiedades del producto vectorial que, aunque muchas de ellas son válidas para cualquier dimensión, se verifican en un espacio vectorial euclídeo tridimensional. Es por lo que repetimos la definición de producto vectorial para $n = 3$, con el fin de justificar mejor las propiedades que vamos a exponer.

5.55' Definición.- *Sean $(E, \cdot)$ un espacio vectorial euclídeo real, $\dim \boldsymbol{E} = 3$ y $\mathcal{B}_0 = \{\vec{\boldsymbol{e}}_1, \vec{\boldsymbol{e}}_2, \vec{\boldsymbol{e}}_3\}$ una base ortonormal, se llama* ***producto vectorial*** *de dos vectores $\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2$, respecto a la orientación determinada por $\mathcal{B}_0$, al vector, denotado por $\vec{\boldsymbol{v}}_1\times\vec{\boldsymbol{v}}_2$, tal que para todo $\vec{\boldsymbol{u}} \in \boldsymbol{E}$, verifica*

$$\vec{\boldsymbol{v}}_1\times\vec{\boldsymbol{v}}_2 \;\cdot\; \vec{\boldsymbol{u}} = \begin{vmatrix} u^1 & u^2 & u^3 \\ v_1^1 & v_1^2 & v_1^3 \\ v_2^1 & v_2^2 & v_2^3 \end{vmatrix}, \qquad (5\text{-}5)$$

donde (u^1, u^2, u^3) y (v_k^1, v_k^2, v_k^n) son las componentes de $\vec{\boldsymbol{u}}$ y $\vec{\boldsymbol{v}}_k$ $(k = 1, 2)$ respecto a la base $\mathcal{B}_0$. Podemos denotar el segundo miembro de la fórmula (5-5) por $\det_{\mathcal{B}_0}(\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2)$.

5.56 Nota.- La notación "$\times$" que hemos adoptado para el producto vectorial está bastante extendida en la literatura matemática. Esta notación se debe a Josiah Willard Gibbs (1839–1903) [(4)], que la utilizó para sus estudiantes de física. Su inconveniente es que puede inducir a una posible confusión con el producto de números reales y el producto cartesiano, pero se entenderá su significado cuando los elementos con los que operamos sean vectores.

Aparece frecuentemente la notación $\vec{u}\wedge\vec{v}$ para el producto vectorial. Esta notación fue iniciada por Cesare Burali-Forti (1861–1931) [(5)] en 1908; Su inconveniente es que puede entrar en conflicto con la notación de producto exterior del álgebra tensorial.

5.57 Proposición[Propiedades del producto vectorial] .- *Sean $\vec{u}$, $\vec{v}$, $\vec{w}$, $\vec{x}$ vectores de un espacio vectorial euclídeo $(E, \cdot)$ de dimensión 3 y $\lambda, \mu \in \mathbb{R}$, todos arbitrario. Se verifica:*

1) $\vec{u}\times\vec{v} = -\vec{v}\times\vec{u}$ *Anticonmutatividad*
2) $(\lambda\vec{u} + \mu\vec{v})\times\vec{w} = \lambda(\vec{u}\times\vec{w}) + \mu(\vec{v}\times\vec{w})$ *Bilinealidad*
3) $\vec{u}\times\vec{v}$ *es ortogonal a* $\vec{u}$ *y a* $\vec{v}$
4) $\vec{u}\times\vec{v} = \vec{o} \Leftrightarrow \vec{u}$ *y a* $\vec{v}$ *son linealmente dependientes.*
5) $\vec{u}\times\vec{v} \neq \vec{o} \Rightarrow \{\vec{u}, \vec{v}, \vec{u}\times\vec{v}\}$ *es una base ordenada con orientación positiva.*
6) $\vec{u}\times\vec{v} \cdot \vec{w} = \vec{u} \cdot \vec{v}\times\vec{w} := [\vec{u}\ \vec{v}\ \vec{w}]$ *Producto mixto*
7) $\vec{u}\times(\vec{v}\times\vec{w}) = (\vec{u}\cdot\vec{w})\vec{v} - (\vec{u}\cdot\vec{v})\vec{w}$. *Doble producto vectorial*
8) $(\vec{u}\times\vec{v})\times\vec{w} + (\vec{v}\times\vec{w})\times\vec{u} + (\vec{w}\times\vec{u})\times\vec{v} = \vec{o}$. *Identidad de Jacobi*
9) $(\vec{u}\times\vec{v})\cdot(\vec{w}\times\vec{x}) = \begin{vmatrix} \vec{u}\cdot\vec{w} & \vec{u}\cdot\vec{x} \\ \vec{v}\cdot\vec{w} & \vec{v}\cdot\vec{x} \end{vmatrix} = (\vec{u}\cdot\vec{w})(\vec{v}\cdot\vec{x}) - (\vec{u}\cdot\vec{x})(\vec{v}\cdot\vec{w})$.
10) $\|\vec{u}\times\vec{v}\|^2 = \|\vec{u}\|^2\|\vec{v}\|^2 - (\vec{u}\cdot\vec{v})^2$. *Módulo del producto vectorial*

Demostración.- Las tres primeras propiedades, una implicación de la 4) y la 6) son consecuencia inmediata de la definición de producto vectorial (5-5) y propiedades de los determinantes.

Para la otra implicación de 4), si $\vec{u}\times\vec{v} = \vec{o}$, todas sus componentes (5-4) son nulas, es decir, el rango de la matriz formada por las dos últimas filas del determinate de la definición es 1; esto quiere decir que las dos últimas filas son dependientes o lo que es lo mismo, $\vec{u}$ y $\vec{v}$ son linealmente dependientes.

[(4)] J. W. Gibbs.- Elements of Vector Análisis. El primer capítulo "Concerning The Algebra of Vectors", lo inicia con definiciones como vector, escalar, y análisis vectorial además de introducir muchos símbolos. En lo referente al producto entre vectores, introdujo el producto directo escrito "$\vec{a}\cdot\vec{b}$", y el "skew product" escrito como "$\vec{a}\times\vec{b}$", que se refiere a los considerados hoy producto escalar y producto vectorial, respectivamente. (http://www.100ciaquimica.net/biograf/cientif/G/gibbs.htm)

[(5)] Cesare Burali-Forti; R. Marcolongo.- Elements de calcul vectoriel avec de nombreuses applications à la geometrie, à la mecanique et à la physique-mathématique, Paris, Hermann, 1910.

5) Se deduce de que $\det(\vec{u}, \vec{v}, \vec{u}\times\vec{v}) = \vec{u}\times\vec{v} \cdot \vec{u}\times\vec{v} = \|\vec{u}\times\vec{v}\| > 0$.

La propiedad 7) puede ser establecida considerando las componentes de los tres vectores y operar en ambos miembros, para comprobar que la igualdad se verifica.

La Identidad de Jacobi, propiedad 8), surge de la propiedad 7).

9) Usando 6), se tiene:
$$(\vec{u}\times\vec{v}) \cdot (\vec{w}\times\vec{x}) = \vec{u} \cdot (\vec{v}\times(\vec{w}\times\vec{x})) = \vec{u} \cdot ((\vec{v} \cdot \vec{x})\vec{w} - (\vec{v} \cdot \vec{w})\vec{x}) =$$
$$(\vec{v} \cdot \vec{x})(\vec{u} \cdot \vec{w}) - (\vec{v} \cdot \vec{w})(\vec{u} \cdot \vec{x}).$$

10) Es consecuencia de la 9), tomando $\vec{w} = \vec{u}$ y $\vec{x} = \vec{v}$.

5.58 Nota.- La propiedad 10) de la Proposición 5.57 "mide" lo que le falta para que la desigualdad de Cauchy–Schwarz (Proposición 5.21) se convierta en una igualdad. Además, nos permite poner la norma del producto vectorial $\vec{u}\times\vec{v}$ en función de las normas de cada uno de ellos y del ángulo que forman. Pues, por ser $\vec{u} \cdot \vec{v} = \|\vec{u}\|\|\vec{v}\| \cos(\widehat{\vec{u}, \vec{v}})$, se obtiene que
$$\|\vec{u}\times\vec{v}\| = \|\vec{u}\|\|\vec{v}\| \operatorname{sen}(\widehat{\vec{u}, \vec{v}}).$$

Las propiedades 1), 2) y 8) de la Proposición 5.57 expresan que el espacio vectorial $I\!R^3$ es una álgebra de Lie [(6)] real respecto al producto vectorial.

5.59 Nota.- Hemos dado la noción de producto vectorial de $n-1$ vectores de un espacio vectorial euclídeo de dimensión n (en particular, de dos vectores en un espacio vectorial euclídeo tridimensional); pero la generalización del producto vectorial de dos vectores a espacios de dimensión mayor que tres no es posible, salvo para el caso de $n = 7$. Este hecho y así como una exposición histórica del origen del producto vectorial y una descripción de las más destacadas y representativas aplicaciones de este producto en Física y Matemáticas, se puede consultar en el artículo:

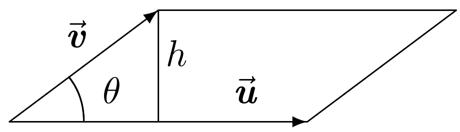
El Producto Vectorial, Juan Francisco González Hernández
http://www.uam.es/personal_pdi/ciencias/fchamizo/realquiler/fich/jfgh.pdf
(Consultada 2-8-2011)

5.60 Ejemplo[Área de un paralelogramo].- *En $I\!R^3$, el módulo del producto vectorial representa el área del paralelogramo cuyos lados son dos vectores con origen en un mismo vértice.*

[(6)] Un álgebra de Lie es un espacio vectorial real dotado de un producto antisimétrico, que en vez de la asociatividad usual verifica una relación llamada identidad de Jacobi; precisando:

Un álgebra de Lie es un espacio vectorial real $\mathfrak{g}$ dotado de una aplicación producto $[\,,]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ tal que:

1. El producto es bilineal.
2. Es antisimétrico: $[y, x] = -[x, y]$.
3. Verifica la identidad de Jacobi: $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$.



$$h = \|\vec{v}\| \operatorname{sen}\theta$$
$$\text{Área}{=}\|\vec{u}\| h = \|\vec{u}\times\vec{v}\|$$

5.61 Ejemplo[Volumen de un paralelepípedo].- *El valor absoluto del producto mixto representa el volumen del paralelepípedo cuyas aristas son tres vectores con origen en un mismo vértice.*

Si $\vec{u}, \vec{v}, \vec{w}$ son vectores en $\mathbb{R}^3$, entonces $|\vec{u}\cdot(\vec{v}\times\vec{w})|$ es el volumen del paralelepípedo definido por los vectores $\vec{u}, \vec{v}, \vec{w}$, se tiene que:

$$|\vec{u}\cdot(\vec{v}\times\vec{w})| = \|\vec{u}\|\,\|\vec{v}\times\vec{w}\|\,|\cos\theta|,$$

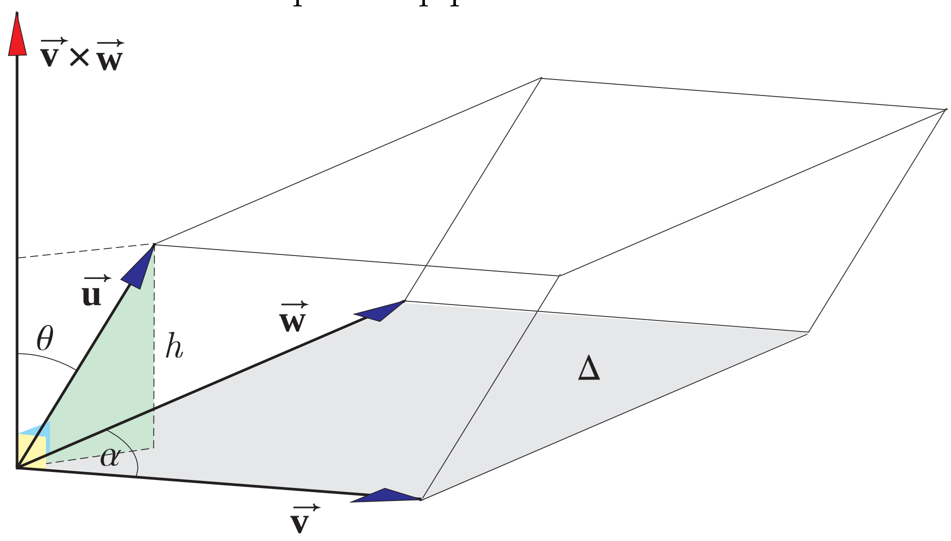
donde θ es el ángulo entre los dos vectores $\vec{u}$ y $\vec{v}\times\vec{w}$.

Por otro lado $\|\vec{v}\times\vec{w}\| = \|\vec{v}\|\,\|\vec{w}\|\,|\operatorname{sen}\alpha|$ corresponde al área del paralelogramo que forman los vectores $\vec{u}, \vec{v}$ y α es el ángulo entre ellos.

Así, reordenando los factores el producto tenemos:

$$\|\vec{u}\cdot(\vec{v}\times\vec{w})\| = (\|\vec{u}\|\,|\cos\theta|)\|\vec{v}\times\vec{w}\| = h\cdot\Delta = V.$$

donde h es la altura del paralelogramo, A es el área del paralelogramo de la base y V es el volumen del paralelepípedo.



5.62 Ejemplo.- *El volumen de un tetraedro cuyos vértices son los puntos $A(3,2,1), B(1,2,4), C(4,0,3)$ y $D(1,1,7)$ es igual a $1/6$ del producto mixto $[\overrightarrow{AB}\,\overrightarrow{AC}\,\overrightarrow{AD}]$, en valor absoluto.*

$\overrightarrow{AB} = (-2,0,3),\ \overrightarrow{AC} = (1,-2,2),\ \overrightarrow{AD} = (-2,-1,6)$
$[\overrightarrow{AB}\,\overrightarrow{AC}\,\overrightarrow{AD}] = 5,\qquad$ Volumen del tetraedro $= 5/6$

5.9 Transformaciones ortogonales

En esta sección consideramos aplicaciones entre espacios vectoriales euclídeos, que además de ser compatibles con su estructura de espacios vectoriales (aplicaciones lineales), conserva los productos escalares definidos en ellos. Estas aplicaciones, llamadas a veces movimientos rígidos, juegan un papel importante en geometría.

5.63 Definición.- *Dados espacios vectoriales euclídeos (E, g) y $(\bar{E}, \bar{g})$, una aplicación lineal $f\colon E \to \bar{E}$ se dice que es **ortogonal** si, para todo $\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \in E$*

$$\bar{g}(f(\vec{\boldsymbol{u}}), f(\vec{\boldsymbol{v}})) = g(\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}). \tag{5-6}$$

*Diremos simplemente que f es una **aplicación ortogonal**.*

5.64 Nota.- Si no hay posibilidad de confusión sobre qué espacios se está trabajando denotamos, como ha sido habitual hasta aquí, la operación de producto escalar en todos los espacios vectoriales euclídeos por "$\cdot$", Así, una aplicación ortogonal $f\colon E \to F$ entre espacios vectoriales euclídeos verifica que:

$$f(\vec{\boldsymbol{u}}) \cdot f(\vec{\boldsymbol{v}}) = \vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}}.$$

Del mismo modo usaremos, si no hay confusión, $\|\ \|$ para las normas asociadas a los productos escalares en los distintos espacios vectoriales euclídeos.

La denominación de aplicación ortogonal, omitiendo el calificativo lineal, está justificada por el hecho de que la condición (5-6) implica que forzosamente la aplicación ha de ser lineal.

En efecto, para demostrar que $f(\lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}) = \lambda f(\vec{\boldsymbol{u}}) + \mu f(\vec{\boldsymbol{v}})$ basta con establecer que:

$$\|f(\lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}) - \lambda f(\vec{\boldsymbol{u}}) - \mu f(\vec{\boldsymbol{v}})\|^2 =$$
$$(f(\lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}) - \lambda f(\vec{\boldsymbol{u}}) - \mu f(\vec{\boldsymbol{v}})) \cdot (f(\lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}) - \lambda f(\vec{\boldsymbol{u}}) - \mu f(\vec{\boldsymbol{v}})) = 0.$$

Efectuando el producto escalar indicado y usando (5-6) resulta:

$$(\lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}) \cdot (\lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}) - \lambda(\lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}) \cdot \vec{\boldsymbol{u}} - \mu(\lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}) \cdot \vec{\boldsymbol{v}} -$$
$$-\lambda\vec{\boldsymbol{u}}(\lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}) \cdot \vec{\boldsymbol{u}} + \lambda^2\vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{u}} + \lambda\mu\vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}} -$$
$$-\mu\vec{\boldsymbol{v}} \cdot (\lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}) + \mu\lambda\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{u}} + \mu\lambda\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{v}} = 0.$$

5.65 Ejemplo.- *Con las estructuras euclídeas canónicas en $\mathbb{R}^2$ y $\mathbb{R}^3$, la aplicación $f\colon \mathbb{R}^2 \to \mathbb{R}^3$*

$$(x, y) \in \mathbb{R}^2 \longrightarrow \left(\frac{x}{\sqrt{2}} + \frac{y}{\sqrt{6}}, \frac{x}{\sqrt{2}} - \frac{y}{\sqrt{6}}, \sqrt{\frac{2}{3}}y\right)$$

es ortogonal.

5.66 Ejemplo.- *Consideremos el espacio euclídeo $\mathbb{R}^3$ con el producto escalar canónico, $(a_0, a_1, a_2) \cdot (b_0, b_1, b_2) = a_0b_0 + a_1b_1 + a_2b_2$, y el espacio euclídeo $\mathcal{C}[0, 2\pi]$ de las funciones continuas definidas en el intervalo $[0, 2\pi]$ con el producto escalar*

$$f \cdot g = \frac{1}{\pi}\int_0^{2\pi} f(t)g(t)dt.$$

Es ortogonal la aplicación

$$F\colon \mathbb{R}^3 \to \mathcal{C}[0, 2\pi], \quad (a_0, a_1, a_2) \mapsto \frac{a_0}{\sqrt{2}} + a_1\cos\theta + a_2\cos 2\theta.$$

$$F(a_0,a_1,a_2)\cdot F(b_0,b_1,b_2)=$$
$$=\frac{1}{\pi}\int_0^{2\pi}\left(\frac{a_0}{\sqrt{2}}+a_1\cos\theta+a_2\cos 2\theta\right)\left(\frac{b_0}{\sqrt{2}}+b_1\cos\theta+b_2\cos 2\theta\right)d\theta=$$
$$=a_0b_0+a_1b_1+a_2b_2=(a_0,a_1,a_2)\cdot(b_0,b_1,b_2).$$

Damos unas propiedades de aplicaciones ortogonales cuya demostración se deducen de forma inmediata de su definición.

5.67 Proposición.- *Dados dos espacios vectoriales euclídeos reales $\boldsymbol{E}$ y $\boldsymbol{F}$, para toda aplicación ortogonal $f\colon\boldsymbol{E}\to\boldsymbol{F}$ se cumple:*

1) $\|f(\vec{\boldsymbol{u}})\|=\|\vec{\boldsymbol{u}}\|,\quad\forall\vec{\boldsymbol{u}}\in\boldsymbol{E}$. *Conserva la norma de vectores*
2) $\vec{\boldsymbol{u}}\perp\vec{\boldsymbol{v}}\Rightarrow f(\vec{\boldsymbol{u}})\perp f(\vec{\boldsymbol{v}}),\quad\forall\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}\in\boldsymbol{E}$. *Conserva la ortogonalidad*
3) *f es inyectiva.*

Demostración.- La propiedad (3) surge de la (1), pues $f(\vec{\boldsymbol{u}})=\vec{o}_F\Rightarrow\vec{\boldsymbol{u}}=\vec{o}_E$. ∎

5.68 Nota.- En la Proposición 5.67 hemos visto que una aplicación ortogonal es una aplicación lineal que conserva la norma de vectores. A modo de recíproco, podemos decir que una aplicación lineal entre espacio vectoriales euclídeos que conserva la norma de vectores es una aplicación ortogonal.

Pues, para todo par de vectores $\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}\in\boldsymbol{E}$:

$$\|f(\vec{\boldsymbol{u}}-\vec{\boldsymbol{v}})\|=\|f(\vec{\boldsymbol{u}})-f(\vec{\boldsymbol{v}})\|=\|\vec{\boldsymbol{u}}-\vec{\boldsymbol{v}}\|\Rightarrow$$
$$(f(\vec{\boldsymbol{u}})-f(\vec{\boldsymbol{v}}))\cdot(f(\vec{\boldsymbol{u}})-f(\vec{\boldsymbol{v}}))=(\vec{\boldsymbol{u}}-\vec{\boldsymbol{v}})\cdot(\vec{\boldsymbol{u}}-\vec{\boldsymbol{v}})\Rightarrow$$
$$\|f(\vec{\boldsymbol{u}})\|^2+\|f(\vec{\boldsymbol{v}})\|-2f(\vec{\boldsymbol{u}})\cdot f(\vec{\boldsymbol{v}})=\|\vec{\boldsymbol{u}}\|^2+\|\vec{\boldsymbol{v}}\|-2\vec{\boldsymbol{u}}\cdot\vec{\boldsymbol{v}}\Rightarrow$$
$$f(\vec{\boldsymbol{u}})\cdot f(\vec{\boldsymbol{v}})=\vec{\boldsymbol{u}}\cdot\vec{\boldsymbol{v}}\Rightarrow f\text{ es ortogonal.}$$

En la situación de que los espacios vectoriales euclídeos $\boldsymbol{E}$ y $\boldsymbol{F}$ tengan la misma dimensión finita una aplicación $f\colon\boldsymbol{E}\to\boldsymbol{F}$ ortogonal es biyectiva. Esto justifica la siguiente definición:

5.69 Definición.- *Sea $\boldsymbol{E}$ un espacio vectorial euclídeo de dimensión finita, se llama **transformación ortogonal** (**isometría lineal** o **movimiento rígido**) sobre $\boldsymbol{E}$ a toda aplicación ortogonal de $\boldsymbol{E}$ en sí mismo.*

5.70 Proposición.- *Sean $\boldsymbol{E}$ un espacio vectorial euclídeo de dimension finita y A la matriz de una aplicación ortogonal $f\colon\boldsymbol{E}\to\boldsymbol{E}$, entonces*

$${}^tAGA=G,$$

donde G es la matriz métrica.

Demostración.- Sea A es la matriz de f respecto a una base $\mathcal{B}=\{\vec{e}_1,\dots,\vec{e}_n\}$ de $\boldsymbol{E}$. Al ser f biyectiva, $\mathcal{B}$' $=\{f(\vec{e}_1),\dots,f(\vec{e}_n)\}$ es una base de $\boldsymbol{E}$ y la matriz cambio de bases de $\mathcal{B}$ a $\mathcal{B}$' es A. Entonces la matriz métrica $(f(\vec{e}_i)\cdot f(\vec{e}_j))_{ij}$ respecto a esta base $\mathcal{B}$' es (ver 5-2) ${}^tAGA$, que coincide, al ser f ortogonal, con la matriz métrica $G=(\vec{e}_i\cdot\vec{e}_j)_{ij}$ respecto a la base $\mathcal{B}$. ∎

5.71 Corolario.- *Sean E un espacio vectorial euclídeo de dimension n y A la matriz de una aplicación ortogonal $f\colon E \longrightarrow E$ en una base ortonormal, entonces A ortogonal:*

$$^{t}AA = I_n,$$

donde I_n es la matriz identidad.

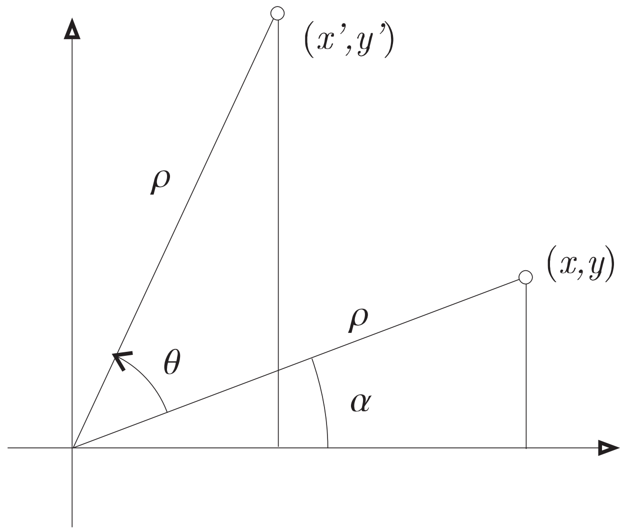
Demostración.- Es consecuencia de que la matriz métrica referida a una base ortonormal es la matriz identidad. ◪

5.72 Ejemplo.- *La aplicación $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ tal que*

$$(x, y) \longmapsto \left(\frac{\sqrt{3}}{2}x + \frac{1}{2}y, -\frac{1}{2}x + \frac{\sqrt{3}}{2}y\right)$$

es ortogonal, respecto al producto escalar canónico en $\mathbb{R}^2$.
Se trata de un giro de 30° al rededor del origen. Cuya expresión general es:

$$\sigma_\theta(x, y) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$



$$x = \rho\cos\alpha \qquad x' = \rho\cos(\alpha+\theta) = \rho\cos\alpha\cos\theta - \rho\operatorname{sen}\alpha\operatorname{sen}\theta = x\cos\theta - y\operatorname{sen}\theta$$
$$y = \rho\operatorname{sen}\alpha \qquad y' = \rho\operatorname{sen}(\alpha+\theta) = \rho\cos\alpha\operatorname{sen}\theta + \rho\operatorname{sen}\alpha\cos\theta = x\operatorname{sen}\theta + y\cos\theta$$

Es inmediato establecer el siguiente resultado:

5.73 Proposición.- *La composición de transformaciones ortogonales es una transformación ortogonal.*
La inversa de una transformación ortogonal también es ortogonal.

◪

Recordar que el determinante de la matriz asociada a una aplicación lineal $f\colon E \longrightarrow E$ es independiente de la elección de la base en E y lo denotamos

por $\det(f)$. También, para toda matriz $A \in \mathcal{M}_{n\times n}(I\!R)$, se tiene que $\det A = \det({}^{\rm t}A)$, y si $B \in \mathcal{M}_{n\times n}(I\!R)$ es otra matiz, $\det(AB) = \det(A)\det(B)$. Entonces, si f es una transformación ortogonal, y si A es su matriz con respecto a una base ortonormal, $A\,{}^{\rm t}A = {}^{\rm t}AA = I$, esto implica que $\det(A)^2 = 1$, esto es, $\det(A) = 1$ o $\det(A) = -1$. Es también claro que las transformaciones ortogonales de un espacio vectorial euclídeo de dimension n forman un grupo (Proposición 5.73) y que las transformaciones ortogonales de determinante $+1$ forman un subgrupo. Todo esto motiva la definición siguiente.

5.74 Definición.- *Dado un espacio vectorial euclídeo $\mathbf{E}$ de dimensión n, el conjunto de las transformaciones ortogonales sobre $\mathbf{E}$ forman un subgrupo del grupo lineal general $GL(\mathbf{E})$ denotado por $O(\mathbf{E})$ o $O(n)$, cuando $\mathbf{E} = I\!R^n$, llamado **grupo ortogonal** (de $\mathbf{E}$) de orden n.*

*Las transformaciones ortogonales tales que $\det(f) = 1$ se llaman **rotaciones** o **isometrías lineales directas (o positivas)** o **transformaciones ortogonales propias** y ellas forman un subgrupo del grupo lineal especial $SL(\mathbf{E})$ (y de $O((E))$, denotado por $SO(\mathbf{E})$ o $SO(n)$, cuando $\mathbf{E} = I\!R^n$, llamado **grupo ortogonal especial** (de $\mathbf{E}$) de orden n.*

*Las transformaciones ortogonales tales que $\det(f) = -1$ se llaman **isometrías lineales inversas (o negativas).***

El siguiente resultado es de gran utilidad para encontrar formas canónicas de la matriz ortogonal asociada a una transformación ortogonal.

5.75 Proposición.- *Sean $\mathbf{E}$ un espacio vectorial euclídeo y $f\colon \mathbf{E} \to \mathbf{E}$ una transformación ortogonal, entonces se tiene:*

a) Si λ es un valor propio de f, entonces $|\lambda| = 1$.

b) Si $\vec{\boldsymbol v}_1$ y $\vec{\boldsymbol v}_2$ son vectores propios correspondientes a valores propios distintos λ_1 y λ_2, entonces $\vec{\boldsymbol v}_1$ y $\vec{\boldsymbol v}_2$ son ortogonales.

c) Si $\mathbf{F}$ es un subespacio vectorial invariante por f, entonces su subespacio ortogonal, $\mathbf{F}^\perp$, es invariante por f.

Demostración.-

a) Si $\vec{\boldsymbol v} \in \mathbf{E} - \{\vec{0}_E\}$ es un vector propio del valor propio λ, $f(\vec{\boldsymbol v}) = \lambda\vec{\boldsymbol v}$, y como f conserva el producto escalar,

$$\vec{\boldsymbol v}\cdot\vec{\boldsymbol v} = f(\vec{\boldsymbol v})\cdot f(\vec{\boldsymbol v}) = \lambda^2\vec{\boldsymbol v}\cdot\vec{\boldsymbol v} \Rightarrow \lambda^2 = 1 \Rightarrow |\lambda| = 1.$$

b) Si $\vec{\boldsymbol v}_1\cdot\vec{\boldsymbol v}_2 \neq 0$, como $\vec{\boldsymbol v}_1\cdot\vec{\boldsymbol v}_2 = f(\vec{\boldsymbol v}_1)\cdot f(\vec{\boldsymbol v}_2) = \lambda_1\lambda_2\vec{\boldsymbol v}_1\cdot\vec{\boldsymbol v}_2$, se tiene que $\lambda_1\lambda_2 = 1$; lo cual es una contradicción ya que λ_1 y λ_2 son distintos y, por a), $|\lambda_1| = |\lambda_2| = 1$.

Se debe verificar que $\vec{\boldsymbol v}_1\cdot\vec{\boldsymbol v}_2 = 0$, es decir, $\vec{\boldsymbol v}_1$, $\vec{\boldsymbol v}_2$ son ortogonales.

c) Sea $\vec{\boldsymbol v} \in \mathbf{F}^\perp$ y veamos que $f(\vec{\boldsymbol v}) \in \mathbf{F}^\perp$.

Para todo $\vec{\boldsymbol u} \in \mathbf{F}$ se tiene que $\vec{\boldsymbol u}\cdot\vec{\boldsymbol v} = 0$ y $\vec{\boldsymbol u}\cdot f(\vec{\boldsymbol v}) = f^{-1}(\vec{\boldsymbol u})\cdot\vec{\boldsymbol v} = 0$, ya que $f^{-1}(\vec{\boldsymbol u}) \in \mathbf{F}$ y f^{-1} es una transformación ortogonal. Así, $f(\vec{\boldsymbol v}) \in \mathbf{F}^\perp$, es decir $f(\mathbf{F}^\perp) \subset \mathbf{F}^\perp$. Y dada la biyectividad de f, $\mathbf{F}^\perp = f(\mathbf{F}^\perp)$. ◾

5.76 Proposición.- *Sean un espacio vectorial euclídeo $\boldsymbol{E}$ de dimensión finita y una transformación ortogonal $f\colon \boldsymbol{E} \to \boldsymbol{E}$, entonces los subespacios* $\mathrm{Ker}(f - 1_{\boldsymbol{E}})$ *y* $\mathrm{Im}(f - 1_{\boldsymbol{E}})$ *son ortogonales y*

$$\boldsymbol{E} = \mathrm{Ker}(f - 1_{\boldsymbol{E}}) \oplus \mathrm{Im}(f - 1_{\boldsymbol{E}})$$

Si $\vec{\boldsymbol{u}} \in \mathrm{Ker}(f - 1_{\boldsymbol{E}})$ y $f(\vec{\boldsymbol{v}}) - v \in \mathrm{Im}(f - 1_{\boldsymbol{E}})$ ($\vec{\boldsymbol{v}} \in \boldsymbol{E}$ arbitrario), al ser f una transformación ortogonal y $f(\vec{\boldsymbol{u}}) = \vec{\boldsymbol{u}}$, se tiene que:

$$\vec{\boldsymbol{u}} \cdot (f(\vec{\boldsymbol{v}}) - \vec{\boldsymbol{v}}) = f(\vec{\boldsymbol{u}}) \cdot f(\vec{\boldsymbol{v}}) - \vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}} = \vec{\boldsymbol{o}}.$$

Así, los subespacios vectoriales $\mathrm{Ker}(f - 1_{\boldsymbol{E}})$ y $\mathrm{Im}(f - 1_{\boldsymbol{E}})$ son ortogonales.
Y como por la Proposición 1.37

$$\dim\ \mathrm{Ker}(f - 1_{\boldsymbol{E}}) + \dim\ \mathrm{Im}(f - 1_{\boldsymbol{E}}) = \dim \boldsymbol{E},$$

tales subespacios son suplementarios. ◼

Enunciamos a continuación un resultado necesario para clasificar movimientos, que en el caso particular del plano y del espacio tridimensional euclídeos son tratados en (§6.4).

5.77 Proposición.- *Sean $\boldsymbol{E}$ un espacio vectorial euclídeo de dimensión n y $f\colon \boldsymbol{E} \to \boldsymbol{E}$ una transformación ortogonal. Entonces, existen una base ortonormal $\{\vec{\boldsymbol{u}}_1, \ldots, \vec{\boldsymbol{u}}_n\}$ de $\boldsymbol{E}$ respecto a la cual la matriz de f es una matriz diagonal por cajas y las cajas son de los tres tipos siguientes:*

$$(1)\,, \quad (-1)\,, \quad \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{con } a^2 + b^2 = 1, b \neq 0.$$

$$\begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& \begin{pmatrix} a & b \\ -b & a \end{pmatrix} && \\ &&&&&&& \ddots & \\ &&&&&&&& \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \end{pmatrix}.$$

TEMA 6
Espacios afines euclídeos

En la geometría afín es posible tratar con relaciones de los vectores y baricentros de puntos, pero no hay manera de expresar el concepto de longitud de un segmento o hablar de ortogonalidad de los vectores. Una estructura euclídea (pág. 87), en el espacio vectorial al cual esta asociado el espacio afín, nos permite hacer frente a las nociones métricas como la ortogonalidad y la longitud (o distancia).

Como la geometría afín es el estudio de las propiedades invariantes bajo aplicaciones biyectivas afines, la geometría euclídea es el estudio de las propiedades invariantes bajo ciertas afinidades llamadas movimientos rígidos. Un movimiento rígido es una aplicación que conservan la distancia entre los puntos. Estas aplicaciones son, de hecho, afines y biyectiva.

6.1 Definición.- *Se dice que un espacio afín $(\mathcal{A}, \boldsymbol{E})$ es* **euclídeo** *si el espacio vectorial real $\boldsymbol{E}$ es euclídeo. Se dirá simplemente que $(\mathcal{A}, \boldsymbol{E})$ es un* **espacio euclídeo**.

Como todo espacio vectorial puede considerarse (Ejemplo 2.4) como un espacio afín asociado a sí mismo, se tiene que un espacio vectorial euclídeo proporciona un primer ejemplo de espacio euclídeo.

6.1 Ortogonalidad en un espacio euclídeo

Damos el concepto de ortogonalidad en espacio euclídeos, teniendo presente la ortogonalidad de vectores dada en §5.4.

6.2 Definición.- *Sea $(\mathcal{A}, \boldsymbol{E})$ un espacio euclídeo de dimensión finita, sea $\mathcal{F} \subseteq \mathcal{A}$ una variedad lineal y sea $P \in \mathcal{A}$. El* **complemento ortogonal** *a $\mathcal{F}$ por el punto P es la variedad lineal $\mathcal{F}_P^\perp = P + \boldsymbol{F}$, donde $\boldsymbol{F}$ es el subespacio director de $\mathcal{F}$.*

6.3 Ejemplo.- *El el espacio euclídeo $\mathbb{R}^3$, determinar los complementos ortogonales de la recta $\mathcal{L} = (1,5,4)+\{(1,2,3)\} \subset \mathbb{R}^3$ por el punto $(1,1,2)$ y del plano $\mathcal{P} = \{(x,y,z) \in \mathbb{R}^3 / 3x-2y+z=7\}$ por el punto $(1,0,1)$.*

$$\mathcal{L}^{\perp}_{(1,1,2)} = (1,1,2) + \{(x,y,z) \in \mathbb{R}^3 / (x,y,z)\cdot(1,2,3) = 0\} =$$

$$= (1,1,2)+\{(x,y,z) \in \mathbb{R}^3 / x+2y+3z = 0\} = \{(x,y,z) \in \mathbb{R}^3 / x+2y+3z = 9\}.$$

Sea el plano $\mathcal{P} = \{x,y,z \in \mathbb{R}^3 / 3x-2y+z=7\}$ su complemento ortogonal por el punto $(1,0,1)$ es la recta, cuya dirección es la del vector, perpendicular al plano, $\vec{v} = (3,-2,1)$:

$$(x,y,z) = (1,0,1) + \lambda(3,-2,1); \qquad \frac{x-1}{3} = -\frac{y}{2} = z-1.$$

6.4 Definición.- *Llamamos **referencia rectangular** en un espacio euclídeo $(\mathcal{A},\boldsymbol{E})$ de dimensión finita ($\dim \boldsymbol{E} = n$) a toda referencia cartesiana $\mathcal{R} = \{O;\vec{e}_1,\dots,\vec{e}_n\}$ tal que $\{\vec{e}_1,\dots,\vec{e}_n\}$ es base ortonormal de $(\boldsymbol{E},\cdot)$.*

6.5 Definición.- *Un vector $\vec{v}$ se dice **ortogonal** a una variedad lineal $\mathcal{F}$ (ponemos $\vec{v} \perp \mathcal{F}$) si $\vec{v}$ es ortogonal al subespacio director $\boldsymbol{F}$ de $\mathcal{F}$, es decir,*

$$\vec{v} \perp \mathcal{F} \Leftrightarrow \vec{v} \in \boldsymbol{F}^{\perp}.$$

En un espacio afín $\mathcal{A}$ de dimensión n, todo hiperplano $\mathcal{H}$ tiene dimensión $n-1$. Luego, el subespacio vectorial ortogonal a su subespacio director $\boldsymbol{H}$ tiene dimensión 1; la dirección de éste queda determinado con un vector ortogonal a $\mathcal{H}$, el cual está relacionado con los coeficientes de la ecuación cartesiana del hiperplano. En efecto:

La ecuación de un hiperplano $\mathcal{H}$ (3-3) es:
$$a_1x^1 + \cdots + a_nx^n = \beta,$$
respecto a una referencia que suponemos rectangular.

Un vector arbitrario de $\boldsymbol{H}$ queda determinado por dos puntos $P(p^1,\dots,p^n)$ y $Q(q^1,\dots,q^n)$ de $\mathcal{H}$, los cuales satisfacen a su ecuación. Por lo que, sustituyendo en ella y restando las dos ecuaciones que resultan, se obtiene
$$a_1(p^1-q^1) + \cdots + a_n(p^n-q^n) = 0.$$
Esto significa, como la referencia es rectangular, que el vector $\vec{v}$ de componentes $(a_1,\dots,a_n)$ es ortogonal al $\overrightarrow{PQ}$, de componentes $(p^1-q^1,\dots,p^n-q^n)$. Así, $\vec{v} \in \boldsymbol{H}^{\perp}$. Concluimos:

6.6 Proposición.- *Respecto a una referencia rectangular en espacio euclídeo de dimensión n, si $a_1x^1+\cdots+a_nx^n = b$ es la ecuación de un hiperplano, el vector de componentes $(a_1,\dots,a_n)$ es ortogonal a dicho hiperplano.*

◪

Pasamos ahora a tratar la ortogonalidad de variedades lineales de un espacio afín dando dos definiciones, una ellas en términos de sus dimensiones cuando la dimensión del espacio afín es n.

6.7 Definición.- *Dos variedades lineales $\mathcal{F}$ y $\mathcal{G}$ de $\mathcal{A}$, de subespacios directores F y G, se dicen que son* `ortogonales` *(o* `perpendiculares`*) si $F \subset G^{\perp}$ o $F \supset G^{\perp}$. Esto se representa poniendo $\mathcal{F} \perp \mathcal{G}$.*

6.8 Definición.- *Sean dos variedades lineales $\mathcal{F}$ y $\mathcal{G}$ de un espacio afín $\mathcal{A}$ de dimension finita n, con subespacios directores F y G, respectivamente.*

1. *Si $\dim \mathcal{F} + \dim \mathcal{G} \leq n$, se dice que $\mathcal{F}$ y $\mathcal{G}$ son* `ortogonales` *(o perpendiculares) si sus espacios directores son ortogonales, es decir, $\mathcal{F} \perp \mathcal{G} \Leftrightarrow F \perp G$.*

2. *Si $\dim \mathcal{F} + \dim \mathcal{G} \geq n$, se dice que $\mathcal{F}$ y $\mathcal{G}$ son* `ortogonales` *(o perpendiculares) si los subespacios ortogonales de sus espacios directores son ortogonales, es decir, $\mathcal{F} \perp \mathcal{G} \Leftrightarrow F^{\perp} \perp G^{\perp}$.*

Hagamos unas aclaraciones sobre estas definiciones y su equivalencia

— Cuando $\dim \mathcal{F} + \dim \mathcal{G} = n$, las nociones de los dos apartados de esta definición coinciden, pues entonces $F \perp G \Leftrightarrow F^{\perp} \perp G^{\perp}$.

Ya que si $F \perp G$, ocurre que $F \subset G^{\perp}$ y $G \subset F^{\perp}$. Y como $\dim F = n - \dim G = \dim G^{\perp}$ y $\dim G = n - \dim F = \dim F^{\perp}$, se tiene que $F = G^{\perp}$ y $G = F^{\perp}$; con lo que $F^{\perp} \perp G^{\perp}$.

Razonando de forma similar que establece la otra implicación.

— Si $F \perp G$ entonces, necesariamente, se tiene que $\dim \mathcal{F} + \dim \mathcal{G} \leq n$.

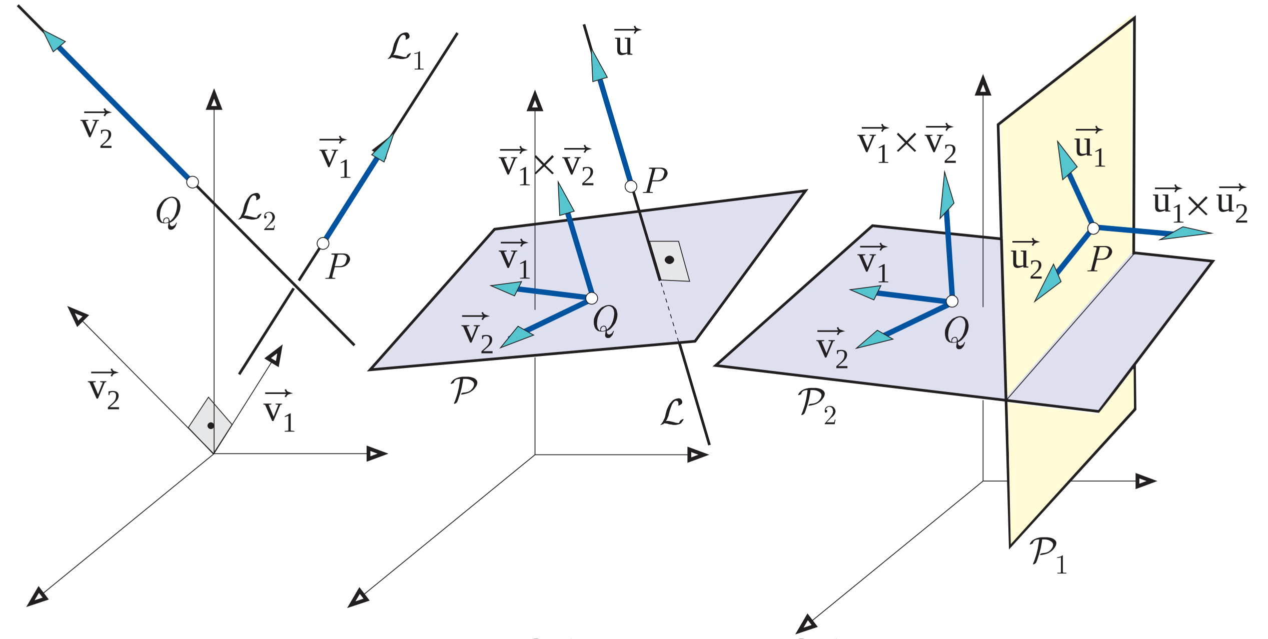
Ya que $F \subset G^{\perp}$ (y $G \subset F^{\perp}$) y, por tanto, $\dim \mathcal{F} = \dim F \leq \dim G^{\perp} = n - \dim G = n - \dim \mathcal{G}$; esto es, $\dim \mathcal{F} + \dim \mathcal{G} \leq n$.

— Así mismo, si $F^{\perp} \perp G^{\perp}$ entonces $\dim \mathcal{F} + \dim \mathcal{G} \geq n$.

Se tiene que $\dim F^{\perp} + \dim G^{\perp} \leq n$; luego, $(n - \dim F) + (n - \dim G) = (n - \dim \mathcal{F}) + (n - \dim \mathcal{G}) \leq n$. Por tanto, $\dim \mathcal{F} + \dim \mathcal{G} \geq n$.

— De la Definición 6.8, en términos de las dimensiones de las variedades lineales, se tiene que $\mathcal{F} \perp \mathcal{G}$ si $F \perp G$ ó $F^{\perp} \perp G^{\perp}$. Cada una de estas condiciones son las que se exigen para que las variedades lineales sean ortogonales en la Definición 6.7. Pues, $F \perp G \Rightarrow F \subset G^{\perp}$ y $F^{\perp} \perp G^{\perp} \Rightarrow G^{\perp} \subset (F^{\perp})^{\perp} = F$.

6.9 Ejemplo.- *En el espacio euclídeo tridimensional $\mathbb{R}^3$, las variedades lineales son las rectas y los planos. Consideremos pares de estas variedades lineales y expresemos las condiciones de perpendicularidad.*



— Sean dos rectas $\mathcal{L}_1 = P + \widetilde{\{\vec{v}_1\}}$ y $\mathcal{L}_2 = Q + \widetilde{\{\vec{v}_2\}}$. Como $\dim \mathcal{L}_1 + \dim \mathcal{L}_2 = 2 < 3 = \dim I\!R^3$, $\mathcal{L}_1 \perp \mathcal{L}_2$ si $\widetilde{\{\vec{v}_1\}} \perp \widetilde{\{\vec{v}_2\}}$. Esto es equivalente a que $\vec{\boldsymbol{v}}_1 \perp \vec{\boldsymbol{v}}_2$.

— Una recta $\mathcal{L} = P + \widetilde{\{\vec{\boldsymbol{u}}\}}$ y un plano $\mathcal{P} = Q + \widetilde{\{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2\}}$ serán ortogonales si $\widetilde{\{\vec{\boldsymbol{u}}\}} \perp \widetilde{\{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2\}}$, o equivalentemente si $\widetilde{\{\vec{\boldsymbol{u}}\}}^{\perp} \perp \widetilde{\{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2\}}^{\perp}$, pues $\dim \mathcal{L} + \dim \mathcal{L} = 3$. Esta condiciones son equivalentes a que $\vec{\boldsymbol{u}} \perp \vec{\boldsymbol{v}}_1$ y $\vec{\boldsymbol{u}} \perp \vec{\boldsymbol{v}}_2$, o también a que $\vec{\boldsymbol{u}} \| (\vec{\boldsymbol{v}}_1 \times \vec{\boldsymbol{v}}_2)$, pues el producto vectorial de dos vectores es ortogonal a ambos.

— Dos planos $\mathcal{P}_1 = P + \widetilde{\{\vec{\boldsymbol{u}}_1, \vec{\boldsymbol{u}}_2\}}$ y $\mathcal{P}_2 = Q + \widetilde{\{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2\}}$ serán ortogonales si $\widetilde{\{\vec{\boldsymbol{u}}_1, \vec{\boldsymbol{u}}_2\}}^{\perp} \perp \widetilde{\{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2\}}^{\perp}$, pues $\dim \mathcal{P}_1 + \dim \mathcal{P}_2 = 4 > 3$. Como el producto vectorial de dos vectores es ortogonal a ambos, esta condición es equivalente a que $(\vec{\boldsymbol{u}}_1 \times \vec{\boldsymbol{u}}_2) \perp (\vec{\boldsymbol{v}}_1 \times \vec{\boldsymbol{v}}_2)$.

6.10 Ejemplo.- *Respecto a una referencia rectangular $\mathcal{R}$ en el espacio euclídeo $I\!R^4$ consideremos el plano $\mathcal{P}$ y el hiperplano $\mathcal{H}$:*

$$\mathcal{P} : \begin{cases} x + z = 2 \\ y + z + t = 1 \end{cases} \qquad \mathcal{H} : \; x - z + t = 3$$

Plano e hiperplano son ortogonales. Las rectas que pasan por el punto $P(0,1,2,3)$ y son perpendiculares a $\mathcal{P}$ están en el plano:

$$y - t = -2, \qquad x + y - z = -1.$$

Dos vectores independientes en el plano por el origen, que es el subespacio director $\boldsymbol{F}$ de $\mathcal{P}$, son $\vec{\boldsymbol{u}}_1 = (1, 0, -1, 1)$ y $\vec{\boldsymbol{u}}_2 = (0, 1, 0, -1)$.

El subespacio ortogonal $\boldsymbol{F}^{\perp}$ a $\boldsymbol{F}$ consta de los vectores $\vec{\boldsymbol{v}} = (v^1, v^2, v^3, v^4)$ tales que $\vec{\boldsymbol{v}} \perp \vec{\boldsymbol{u}}_1$ y $\vec{\boldsymbol{v}} \perp \vec{\boldsymbol{u}}_2$:

$$v^1 - v^3 + v^4 = 0, \qquad v^2 - v^4 = 0.$$

Con lo que dos vectores independientes de $\boldsymbol{F}^{\perp}$ son $\vec{\boldsymbol{v}}_1 = (1, 0, 1, 0)$ y $\vec{\boldsymbol{v}}_2 = (0, 1, 1, 1)$.

Como $\dim \mathcal{P} + \dim \mathcal{H} = 2 + 3 = 5 > 4$ y teniendo en cuenta que el vector $\vec{\boldsymbol{w}} = (1, 0, -1, 1)$ es ortogonal al hiperplano $\mathcal{H}$ (Proposición 6.6), sabemos que $\mathcal{H} \perp \mathcal{P} \Leftrightarrow \vec{\boldsymbol{w}} \perp \vec{\boldsymbol{v}}_1$ y $\vec{\boldsymbol{w}} \perp \vec{\boldsymbol{v}}_2$. Condiciones éstas que se comprueban fácilmente.

Las rectas que pasan por $P(0, 1, 2, 3)$ y son perpendiculares a $\mathcal{P}$ tienen la dirección dada por un vector en $\boldsymbol{F}^{\perp}$. Así, los puntos del plano que las contiene se expresan paramétricamente en la forma:

$$(x, y, z, t) = (0, 1, 2, 3) + \lambda(1, 0, 1, 0) + \mu(0, 1, 1, 1).$$

Eliminando λ y μ, se obtiene las ecuaciones cartesianas:

$$y - t = -2, \qquad x + y - z = -1.$$

Obsérvese que el hiperplano paralelo a $\mathcal{H}$ que pasa por $(0, 1, 2, 3)$, contiene a este plano. Dicho hiperplano tiene por ecuación $x - z + t = b$; por lo que $b = 1$.

Proyección y simetría ortogonales respecto a una variedad lineal

6.11 Proposición.- *Dado un variedad lineal $\mathcal{F} = P + \boldsymbol{F}$ de un espacio afín $(\mathcal{A}, \boldsymbol{E})$ de dimensión finita, para cada punto $X \in \mathcal{A}$ existe un único punto $Y \in \mathcal{F}$ tal que $\overrightarrow{XY} \in \boldsymbol{F}^{\perp}$. Al punto Y le llamamos **proyección ortogonal** de X sobre $\mathcal{F}$ y escribimos $Y = p_{\mathcal{F}}(X)$.*

Demostración.- Tomemos el vector $\overrightarrow{PX} \in \boldsymbol{E}$ y consideremos la proyección ortogonal sobre el subespacio $\boldsymbol{F}$, $p_{\boldsymbol{F}}(\overrightarrow{PX}) \in \boldsymbol{F}$; entonces, si $Y = P + p_{\boldsymbol{F}}(\overrightarrow{PX})$ ($\overrightarrow{PY} = p_{\boldsymbol{F}}(\overrightarrow{PX})$) y como se tiene que:

$$\overrightarrow{PX} = \overrightarrow{PY} + \overrightarrow{YX}, \qquad \overrightarrow{PX} = p_{\boldsymbol{F}}(\overrightarrow{PX}) + p_{\boldsymbol{F}^{\perp}}(\overrightarrow{PX}),$$

resulta que $\overrightarrow{YX} = p_{\boldsymbol{F}^{\perp}}(\overrightarrow{PX}) \in \boldsymbol{F}^{\perp}$. Y como la segunda descomposición es única, al ser $\boldsymbol{E} = \boldsymbol{F} \oplus \boldsymbol{F}^{\perp}$, El punto Y es único. ◪

Este resultado nos permite determinar, de una forma practica, la proyección ortogonal $p_{\mathcal{F}}(X)$ de un punto $X \in \mathcal{A}$ sobre una variedad lineal $\mathcal{F}$, ya que

$$\{p_{\mathcal{F}}(X)\} = \mathcal{F} \cap \mathcal{F}_X^{\perp},$$

siendo $\mathcal{F}_X^{\perp} = X + F^{\perp}$. Pues $\mathcal{F} \cap \mathcal{F}_X^{\perp} = p_{\mathcal{F}}(X) + (\boldsymbol{F} \cap \boldsymbol{F}^{\perp})$ y $\boldsymbol{F} \cap \boldsymbol{F}^{\perp} = \{\vec{\boldsymbol{0}}_E\}$.

6.12 Ejemplo.- *En el espacio euclídeo $\mathbb{R}^3$, proyección ortogonal Y del punto $X(\xi, \eta, \zeta)$ sobre la recta que pasa por los puntos $P(0, 1, 1)$ y $Q(2, 0, 2)$.*

Para determinar la proyección Y de X sobre la recta dada $\mathcal{F} = P + \{\overrightarrow{PQ}\}$, es fácil determinar $\mathcal{F}_X^{\perp}$ (que en este caso es un plano), puesto que el vector director de la recta $\overrightarrow{PQ} = (2, -1, 1)$ es perpendicular al plano. Así (Proposición 6.6):

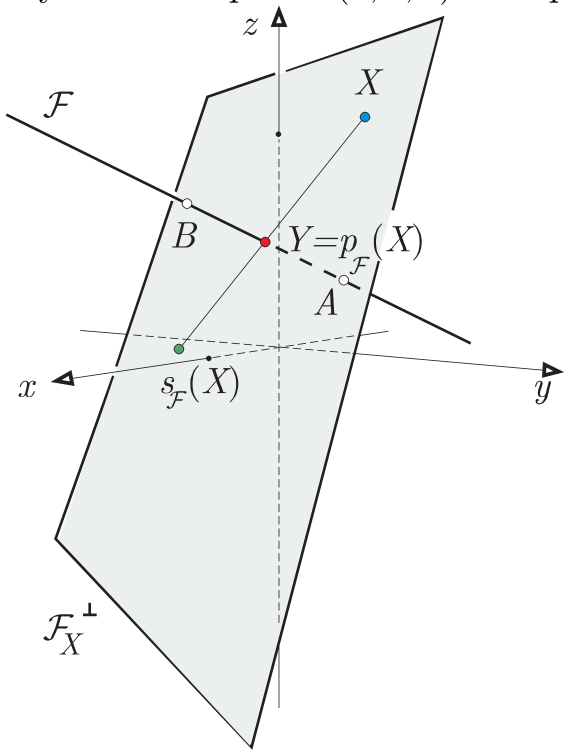
$$\mathcal{F}_X^{\perp} : \ 2x - y + z = d,$$

y como $X \in \mathcal{F}_X^{\perp}$, $d = 2\xi - \eta + \zeta$.

Para determinar la intersección de plano y recta, usamos la ecuación paramétrica de ésta:

$$(x, y, z) = (0, 1, 1) + t(2, -1, 1).$$

Sustituyendo en la ecuación del plano obtenemos el punto
$$\left(\frac{1}{3}(\zeta-\eta+2\xi), \frac{1}{6}(-\zeta+\eta-2\xi+6), \frac{1}{6}(\zeta-\eta+2\xi+6)\right).$$
En particular, la proyección del punto $(1,2,3)$ es el punto $(1,1/2,3/2)$.



Aprovechando la proyección ortogonal sobre una variedad lineal introducimos aquí la simetría ortogonal como un caso particular de las simetrías en la dirección de una variedades lineal cualquiera (§ 4.6).

6.13 Definición.- *Sea $\mathcal{A}$ un espacio afín de dimensión finita asociado a un espacio vectorial $\mathbf{E}$; dada una variedad lineal $\mathfrak{F}$ con subespacio director $\mathbf{F}$ y dado un punto $X \in \mathcal{A}$, se llama **simétrico ortogonal** de X respecto a $\mathfrak{F}$ al punto $s_{\mathfrak{F}}(X)$, tal que*
$$s_{\mathfrak{F}}(X) = X + 2\overrightarrow{Xp_{\mathfrak{F}}(X)},$$
siendo $p_{\mathfrak{F}}$ la proyección ortogonal de X sobre $\mathfrak{F}$.
*Se dice que $s_{\mathfrak{F}}$ es la **simetría ortogonal** respecto a $\mathfrak{F}$.*

Una simetría ortogonal es una aplicación afín, pues podemos escribir:
$$s_{\mathfrak{F}}(X) = X + 2\overrightarrow{Xp_{\mathfrak{F}}(X)} = P + \overrightarrow{PX} + 2\overrightarrow{X(P{+}p_{\mathbf{F}}(\overrightarrow{PX}))} =$$
$$= P + \overrightarrow{PX} + 2\overrightarrow{XP} + 2p_{\mathbf{F}}(\overrightarrow{PX}) = P + (2p_{\mathbf{F}} - 1_{\mathbf{E}})(\overrightarrow{PX}) = P + s_{\mathbf{F}}(\overrightarrow{PX}).$$
Con lo que la simetría $s_{\mathbf{F}}$ es su aplicación lineal asociada.

En el Ejemplo 6.12, la simetría ortogonal respecto a la recta PQ es:
$$s_{\mathfrak{F}}(x,y,z) = \frac{1}{3}\big(x-2y+2z, -2x-2y-z+6, 2x-y-2z+6\big).$$
Y, en particular, el simétrico de $(1,2,3)$ es el punto de coordenadas $(1,-1,0)$.

6.14 Ejemplo.- *Sean en el espacio euclídeo $\mathbb{R}^4$, respecto a una referencia rectangular, el punto $P(1,1,0,-1)$ y la recta $\mathfrak{L}$, dada por sus ecuaciones paramétricas:*
$$x=\lambda, \quad y=2\lambda-1, \quad z=-1, \quad t=-\lambda.$$

La proyección ortogonal del punto P sobre la recta $\mathcal{L}$ es el punto $p_{\mathcal{L}}(P) = (1, 1, -1, -1)$.

Vamos a determinar la proyección ortogonal, $Y(\lambda, 2\lambda - 1, -1, -\lambda)$, de un punto genérico $X(\xi, \eta, \zeta, \tau) \in I\!R^4$ sobre $\mathcal{L}$.

La recta XY ha de ser perpendicular a $\mathcal{L}$; luego, han de ser perpendiculares sus direcciones, dadas por los vectores $\overrightarrow{XY} = (2\lambda - \xi, \lambda - \eta - 1, -\zeta - 1, -\lambda - \tau)$ y $\vec{\boldsymbol{v}} = (1, 2, 0, -1)$, respectivamente. O sea:

$$\overrightarrow{XY} \perp \vec{\boldsymbol{v}} \Leftrightarrow (2\lambda - \xi, \lambda - \eta - 1, -\zeta - 1, -\lambda - \tau) \cdot (1, 2, 0, -1) = 0.$$

Luego, $\lambda = \frac{1}{6}(2\eta + \xi - \tau + 2)$ y la proyección de $X(\xi, \eta, \zeta, \tau)$ sobre $\mathcal{L}$ es:

$$p_{\mathcal{F}}(\xi, \eta, \zeta, \tau) = \left(\frac{1}{6}(2\eta + \xi - \tau + 2), \frac{1}{3}(2\eta + \xi - \tau - 1), -1, \frac{1}{6}(-2\eta - \xi + \tau - 2)\right)$$

En particular, si $X = P(1, 1, 0, -1)$ su proyección es $(1, 1, -1, 1)$

Comprobemos que este punto es la intersección: $\mathcal{L} \cap \mathcal{L}_X^{\perp}$.

La variedad lineal ortogonal a $\mathcal{L}$ que pasa por $X(\xi, \eta, \zeta, \tau)$ es un hiperplano, cuyo subespacio director es perpendicular al vector $\vec{\boldsymbol{v}} = (1, 2, 0, -1)$; luego su ecuación será:

$$x + 2y - t = d, \text{ con } d = \xi + 2\eta - \tau.$$

Sustituyendo en esta ecuación un punto genérico de la recta, $x = \lambda, y = 2\lambda - 1, z = -1, t = -\lambda$, se obtiene

$$\lambda + 2(2\lambda - 1) - (-\lambda) = \xi + 2\eta - \tau \Rightarrow \lambda = \frac{1}{6}(2\eta + \xi - \tau + 2).$$

6.2 Distancia en un espacio euclídeo

La existencia de un producto escalar en el espacio vectorial permite hablar de distancia entre puntos de un espacio euclídeo.

6.15 Definición.- *Se llama **distancia** (euclídea) entre dos puntos P y Q de un espacio euclídeo $\mathcal{A}$ al número*

$$d(P, Q) = \|\overrightarrow{PQ}\|.$$

La siguiente proposición enuncia las propiedades de la distancia, las cuales se deducen inmediatamente de las propiedades (Proposición 5.21) de la norma asociada al producto escalar dado en E.

6.16 Proposición.- *La aplicación distancia $d: \mathcal{A} \times \mathcal{A} \to I\!R$, definida por $d(X, Y) = \|\overrightarrow{XY}\|$, verifica las siguientes propiedades, $\forall P, Q, R \in \mathcal{A}$:*

1. $d(P, Q) = d(Q, P)$.
2. $d(P, Q) \geq 0$ *y* $d(P, Q) = 0 \Leftrightarrow P = Q$.
3. $d(P, R) \leq d(P, Q) + d(Q, R)$ *(Desigualdad triangular).*

$\blacksquare$

Como consecuencia de estas propiedades, se verifica $\forall P, Q, R \in \mathcal{A}$:
$$d(P,Q) \geq |d(P,R) - d(Q,R)|.$$
Pues, por la desigualdad triangular,
$$d(P,R) \leq d(P,Q) + d(Q,R), \qquad d(Q,R) \leq d(Q,P) + d(P,R).$$
De donde resultan las siguientes relaciones, de las cuales se obtiene el resultado enunciado:
$$d(P,Q) \geq d(P,R) - d(Q,R),$$
$$d(Q,P) \geq d(Q,R) - d(P,R) = -(d(P,R) - d(Q,R)).$$

La distancia euclídea se conserva por traslaciones, es decir, si P, Q son puntos de un espacio euclídeo $(\mathcal{A}, \boldsymbol{E})$ y $\vec{\boldsymbol{u}} \in \boldsymbol{E}$:
$$d(P + \vec{\boldsymbol{u}}, Q + \vec{\boldsymbol{u}}) = d(P,Q), \quad \text{o} \quad d(t_{\vec{\boldsymbol{u}}}(P), t_{\vec{\boldsymbol{u}}}(Q)) = d(P,Q).$$
En efecto, es consecuencia de que $\overrightarrow{(P + \vec{\boldsymbol{u}})(Q + \vec{\boldsymbol{u}})} = \overrightarrow{PQ}$.

Esta propiedad nos da un ejemplo de isometrías (§6.3), que son afinidades que conservan la distancia.

Determinemos la expresión de la distancia entre puntos, en términos de las coordenadas de éstos respecto a una referencia rectangular.

6.17 Proposición.- *Sean dos puntos $P(p^1, \ldots, p^n)$ y $Q(q^1, \ldots, q^n)$, dados por sus coordenadas cartesianas respecto a una referencia rectangular $\mathcal{R} = \{O; \vec{\boldsymbol{e}}_1, \ldots, \vec{\boldsymbol{e}}_n\}$, la distancia entre ellos viene dada por:*
$$d(P,Q) = \sqrt{(p^1 - q^1)^2 + \cdots + (p^n - q^n)^2}.$$

Demostración.- Basta tener en cuenta las componentes $(p^1 - q^1, \ldots, p^n - q^n)$ del vector $\overrightarrow{PQ}$ en la referencia $\mathcal{R}$ y que ésta es rectangular. ◪

6.18 Proposición[Teorema de Pitágoras].- *En un espacio euclídeo, sean A, B, C son tres puntos no alineados tales que $\overrightarrow{AB} \perp \overrightarrow{AC}$ son ortogonales, entonces:*
$$d(B,C)^2 = d(A,B)^2 + d(A,C)^2.$$

Demostración.-
$$\overrightarrow{BC} \cdot \overrightarrow{BC} = (\overrightarrow{AC} - \overrightarrow{AB}) \cdot (\overrightarrow{AC} - \overrightarrow{AB}) = \overrightarrow{AC} \cdot \overrightarrow{AC} + \overrightarrow{AB} \cdot \overrightarrow{AB} - 2\overrightarrow{AB} \cdot \overrightarrow{AC} =$$
$$\overrightarrow{AC} \cdot \overrightarrow{AC} + \overrightarrow{AB} \cdot \overrightarrow{AB}.$$
◪

6.19 Ejemplo.- *Sea $P, Q \in \mathcal{A}$ son dos puntos distintos, el lugar geométrico de los puntos que equidistan de P y Q es el hiperplano*
$$\mathcal{H} = M + \left(\widetilde{\{\overrightarrow{PQ}\}}\right)^{\perp},$$
donde M es el punto medio del segmento $\overline{PQ}$. En el caso particular en que $\dim \mathcal{A} = 2$, tal lugar geométrico es la recta mediatriz del segmento $\overline{PQ}$.

En efecto, sea $X \in \mathcal{A}$, como M es el punto medio de $\overline{PQ}$, se tienen las relaciones $\|\overrightarrow{MP}\| = \|\overrightarrow{MQ}\|$ y $\overrightarrow{QM} = -\overrightarrow{PM}$, que usamos a continuación:

$$d(P,X) = d(Q,X) \Leftrightarrow \|\overrightarrow{PX}\|^2 = \|\overrightarrow{QX}\|^2 \Leftrightarrow$$
$$\Leftrightarrow (\overrightarrow{PM} - \overrightarrow{MX}) \cdot (\overrightarrow{PM} - \overrightarrow{MX}) = (\overrightarrow{QM} - \overrightarrow{MX}) \cdot (\overrightarrow{QM} - \overrightarrow{MX}) \Leftrightarrow$$
$$\Leftrightarrow \overrightarrow{PM} \cdot \overrightarrow{MX} = \overrightarrow{QM} \cdot \overrightarrow{MX} \Leftrightarrow \overrightarrow{PM} \cdot \overrightarrow{MX} = -\overrightarrow{PM} \cdot \overrightarrow{MX} \Leftrightarrow \overrightarrow{PM} \cdot \overrightarrow{MX} = 0.$$

Esta última relación es la ecuación del hiperplano que pasa por M y es perpendicular $\overrightarrow{PM}$.

Distancia entre variedades lineales

La noción de proyección ortogonal nos permite encontrar la mínima distancia de un punto a una variedad. Lo cual utilizaremos para hallar la distancia entre variedades lineales, sólo en unos casos particulares concretos, remitiendo, para un estudio sobre distancias entre variedades afines cualesquiera, a [2, Lección 32]

Empezamos tratando la distancia de un punto a una variedad lineal de un espacio euclídeo de dimensión finita.

6.20 Definición.- *Dado un espacio euclídeo $\mathcal{A}$, sean una variedad lineal no vacía $\mathcal{F}$ de $\mathcal{A}$ y un punto P un punto de $\mathcal{A}$. Definimos la* **distancia** *de P a $\mathcal{F}$ como*

$$d(P,\mathcal{F}) = \inf\{\|\overrightarrow{PX}\| / X \in \mathcal{F}\}.$$

Como $\{\|\overrightarrow{PX}\| / X \in \mathcal{F}\} \subseteq \mathbb{R}^+$ es un subconjunto de número reales no negativos, tiene un ínfimo, el cual pertenece a dicho conjunto (y por tanto, es mínimo) si sólo si el subespacio director de $\mathcal{F}$ admite un suplementario ortogonal. Lo cual ocurre, por ejemplo, cuando el espacio afín es de dimensión finita. Entonces, la distancia de un punto a la variedad es la distancia de dicho punto a su proyección ortogonal sobre la variedad lineal.

6.21 Proposición.- *Dado un punto Q y una variedad lineal $\mathcal{F} = P + \boldsymbol{F}$ de un espacio euclídeo $(\mathcal{A}, \boldsymbol{E})$ de dimensión n, se verifica que $d(Q,\mathcal{F}) = d(Q, p_{\mathcal{F}}(Q))$. Además la proyección ortogonal $p_{\mathcal{F}}(Q)$ es el único punto de $\mathcal{F}$ que verifica esta igualdad.*

Demostración.- Sea $R = p_{\mathcal{F}}(Q)$, entonces $\overrightarrow{PR} \in \boldsymbol{F}$ y $\overrightarrow{RQ} \in \boldsymbol{F}^{\perp}$, es decir, $\overrightarrow{PR} \perp \overrightarrow{RQ}$. Aplicando el Teorema de Pitágoras (Proposición 6.18),

$$d(P,Q)^2 = d(P,R)^2 + d(R,Q)^2.$$

Como podemos sustituir el punto P por cualquier otro punto de $\mathcal{F}$, esta igualdad se satisface para cualquier punto P de $\mathcal{F}$ y, por tanto, si $P \neq R$, se tiene $d(P,R) > 0$ y $d(P,Q)^2 > d(R,Q)^2$. Si $P = R$, $d(P,Q)^2 = d(R,Q)^2$.

Se concluye que para todo punto $P \in \mathcal{F}$, distinto de Q, $d(P,Q) > d(R,Q)$; es decir, $d(Q,\mathcal{F}) = d(R,Q)$. ◼

6.22 Ejemplo.- *Con los datos dados en el Ejemplo 6.10, la distancia del punto $P(0,1,2,3)$ al plano $\mathcal{P}$ $(x+z=2,\ y+z+t=1)$ es $\sqrt{10}$.*

Tomando el punto $P_0(2,1,0,0) \in \mathcal{P}$, para determinar la proyección ortogonal $p_{\mathcal{F}}(P)$ de P sobre $\mathcal{P}$, debemos calcular la proyección del vector $\overrightarrow{P_0P}$ sobre $\boldsymbol{F}$ (subespacio director de $\mathcal{P}$), para lo que ponemos:

$$\overrightarrow{P_0P} = \lambda_1\vec{\boldsymbol{u}}_1 + \lambda_2\vec{\boldsymbol{u}}_2 + \mu_1\vec{\boldsymbol{v}}_1 + \mu_2\vec{\boldsymbol{v}}_2.$$

Siendo $\overrightarrow{P_0P} = (-2,0,2,3)$; $\vec{\boldsymbol{u}}_1 = (1,0,-1,1)$ y $\vec{\boldsymbol{u}}_2 = (0,1,0,-1)$ dos vectores que generan $\boldsymbol{F}$; y $\vec{\boldsymbol{v}}_1 = (1,0,1,0)$ y $\vec{\boldsymbol{v}}_2 = (0,1,1,1)$ los que generan $\boldsymbol{F}^{\perp}$.

Los valores que se obtienen son $\lambda_1 = -1, \lambda_2 = -2, \mu_1 = -1, \mu_2 = 2$. Luego:

$$p_{\boldsymbol{F}}(\overrightarrow{P_0P}) = (-1,-2,1,1), \qquad p_{\boldsymbol{F}^{\perp}}(\overrightarrow{P_0P}) = (-1,2,1,2).$$

La distancia $d(P,\mathcal{P}) = \|p_{\boldsymbol{F}^{\perp}}(\overrightarrow{P_0P})\| = \sqrt{10}$.

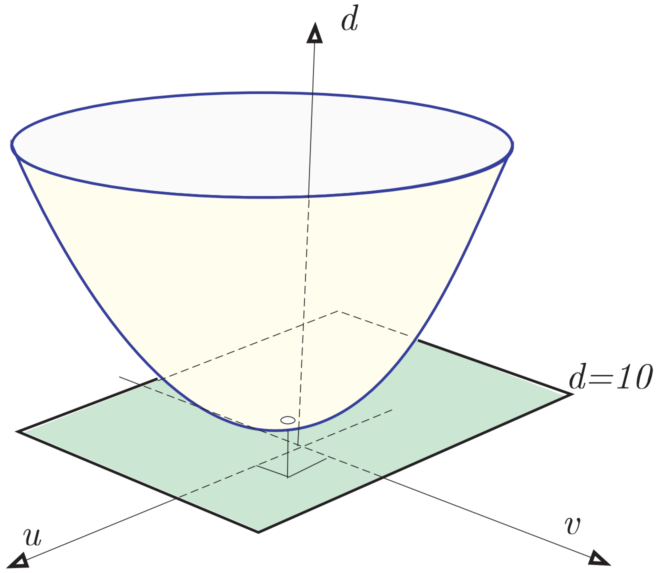
El punto de $\mathcal{P}$ más cercano de P es su proyección ortogonal $p_{\mathcal{P}}(P)$ sobre $\mathcal{P}$:

$$p_{\mathcal{P}}(P) = P_0 + p_{\boldsymbol{F}}(P) = P - p_{\boldsymbol{F}^{\perp}}(P) = (1,-1,1,1).$$

6.23 Nota.- Podemos proceder para hallar la distancia de P al plano $\mathcal{P}$, de este Ejemplo, determinando la distancia de $P(0,1,2,3)$ a un punto genérico $X(u,v,-u+2,u-v-1)$:

$$d(P,X)^2 = f(u,v) = 3u^2 + 2v^2 - 2uv - 8u + 6v + 17.$$

La función f alcanza un mínimo en $(u_0,v_0) = (1,-1)$ y este mínimo es $f(1,-1) = 10$. El punto X_0 del plano $\mathcal{P}$ para el cual la distancia es mínima es entonces $X_0(1,-1,1,1)$.



En el espacio coordenado u, v, d (las escalas en los ejes son diferentes), $d = 3u^2 + 2v^2 - 2uv - 8u + 6v + 17$ es un paraboloide elíptico y $d = 10$ es el plano tangente en $(1,-1,10)$.

En el caso particular de un hiperplano podemos expresar la distancia de un punto al hiperplano en función de los coeficientes de su ecuación respecto a una referencia rectangular:

6.24 Proposición.- *Si $a_1x^1 + \cdots + a_nx^n + b = 0$ es la ecuación de un hiperplano $\mathcal{H}$ y $(p^1,\ldots,p^n)$ son las coordenadas de un punto P respecto a una referencia rectangular en un espacio euclídeo, la distancia de P a $\mathcal{H}$ viene dada por*

$$d(p,\mathcal{H}) = \frac{|a_1p^1 + \cdots + a_np^n + b|}{\sqrt{a_1^2 + \cdots + a_n^2}}.$$

Demostración.- Si $P \in \mathcal{H}$ la fórmula es cierta. Sea $P \notin \mathcal{H}$. La ecuación de $\mathcal{H}$ se puede escribir como:

$$\frac{a_1 x^1 + \cdots + a_n x^n + b}{\sqrt{a_1^2 + \cdots + a_n^2}} = a'_1 x^1 + \cdots + a'_n x^n + b' = 0.$$

El vector $\vec{\boldsymbol{u}} = (a'_1, \ldots, a'_n)$ es unitario y (Proposición 6.6) $\vec{\boldsymbol{u}} \in \boldsymbol{H}^{\perp}$, es decir, pertenece al suplementario ortogonal (unidimensional) del subespacio director $\boldsymbol{H}$ de $\mathcal{H}$.

Se sabe que $d(P, \mathcal{H}) = d(P, P_0)$, con $P_0 \in \mathcal{H}$ y $\overrightarrow{P_0 P} \in \boldsymbol{H}^{\perp}$. Luego, $\overrightarrow{P_0 P} = \lambda \vec{\boldsymbol{u}}$ ($\lambda \in I\!R$), con lo que $d(P, \mathcal{H}) = |\lambda|$. Usando coordenadas, podemos poner:

$$(p^1 - p_0^1, \ldots, p^n - p_0^n) \cdot (a'_1, \ldots, a'_n) = \lambda \Rightarrow p^1 a'_1 + \cdots + p^n a'_n + b' = \lambda.$$

Ahora vamos a definir la distancia entre dos variedades lineales no vacías, aunque no existen dos puntos únicos (uno en cada una de ellas) cuya distancia entre ellos sea la distancia entre las variedades lineales (por ejemplo, dos variedades paralelas). También se trata la variedad lineal perpendicular a dos variedades lineales disjuntas.

6.25 Definición.- *Dado un espacio euclídeo $\mathcal{A}$, sean dos variedades lineales no vacía y disjuntas, $\mathcal{F}_1$ y $\mathcal{F}_2$ de $\mathcal{A}$. Definimos la* ***distancia*** *de $\mathcal{F}_1$ a $\mathcal{F}_2$ como*

$$d(\mathcal{F}_1, \mathcal{F}_2) = inf\,\{d(P_1, P_2) / P_1 \in \mathcal{F}_1, P_2 \in \mathcal{F}_2\}.$$

Para el caso de espacios euclídeos de dimensión finita, existe dos puntos, uno en cada variedad lineal, tales que la distancias entre ellos coincide con la distancia entre las variedades lineales, como se pondrá de manifiesto a continuación.

6.26 Proposición.- *Sean $\mathcal{F}_1$ y $\mathcal{F}_2$ variedades no vacías y disjuntas en un espacio euclídeo $(\mathcal{A}, \boldsymbol{E})$ de dimension n. Existe una variedad lineal $\mathcal{F}$ (de subespacio director $\boldsymbol{F}$), llamada* ***variedad lineal perpendicular común*** *a $\mathcal{F}_1$ y $\mathcal{F}_2$, tal que:*

1) $\dim \mathcal{F} \geq 1$.
2) $\mathcal{F} \perp \mathcal{F}_i \quad (i = 1, 2)$.
3) $\mathcal{F} \cap \mathcal{F}_i = \{P_i\} \quad (i = 1, 2)$. *A los puntos P_1 y P_2 se les denomina* ***pies*** *de la variedad lineal perpendicular común.*
4) *Si $\mathcal{F}'$ es otra variedad lineal que verifica 1), 2) y 3), de subespacio director $\boldsymbol{F}'$, se verifica $\boldsymbol{F}' \subset \boldsymbol{F}$.*
5) $d(\mathcal{F}_1, \mathcal{F}_2) = d(P_1, P_2)$.
6) *$\mathcal{F}$ es única $\Leftrightarrow$ $\mathcal{F}_1$ y $\mathcal{F}_2$ no son paralelas.*

Demostración.- Como $\mathcal{F}_1 \cap \mathcal{F}_2 = \varnothing$ se tiene, por la Proposición 2.23, que $\boldsymbol{F}_1 + \boldsymbol{F}_2 \subsetneq \boldsymbol{E}$. Tomemos el subespacio vectorial de $\boldsymbol{E}$:

$$\boldsymbol{F} = (\boldsymbol{F}_1 + \boldsymbol{F}_2)^{\perp} = \boldsymbol{F}_1^{\perp} \cap \boldsymbol{F}_2^{\perp}.$$

Se verifica que $\dim \boldsymbol{F} = n - \dim(\boldsymbol{F}_1 + \boldsymbol{F}_2) \geq 1$ y además $\boldsymbol{E} = (\boldsymbol{F}_1 + \boldsymbol{F}_2) \oplus \boldsymbol{F}$.

Consideremos puntos $Q_1 \in \mathcal{F}_1$ y $Q_2 \in \mathcal{F}_2$, se tiene la descomposición única del vector $\overrightarrow{Q_1Q_2} = \vec{\boldsymbol{v}} + \vec{\boldsymbol{u}}$, con $\vec{\boldsymbol{v}} \in \boldsymbol{F}_1 + \boldsymbol{F}_2$ y $\vec{\boldsymbol{u}} \in \boldsymbol{F}$; pongamos $\vec{\boldsymbol{v}} = \vec{\boldsymbol{v}}_1 + \vec{\boldsymbol{v}}_2$, con $\vec{\boldsymbol{v}}_1 \in \boldsymbol{F}_1$ y $\vec{\boldsymbol{v}}_2 \in \boldsymbol{F}_2$ y sean:

$$\boxed{P_1 = Q_1 + \vec{\boldsymbol{v}}_1 \in \mathcal{F}_1, \qquad P_2 = Q_2 - \vec{\boldsymbol{v}}_2 \in \mathcal{F}_2, \qquad \mathcal{F} = P_1 + \boldsymbol{F}.} \tag{6-1}$$

— La variedad $\mathcal{F}$ verifica claramente las dos primeras afirmaciones, 1) y 2).

— Para probar 3), observemos que $P_1 \in \mathcal{F} \cup \mathcal{F}_1$, y como

$$P_2 = Q_2 - \vec{\boldsymbol{v}}_2 = Q_1 + \overrightarrow{Q_1Q_2} - \vec{\boldsymbol{v}}_2 = Q_1 + \vec{\boldsymbol{v}}_1 + \vec{\boldsymbol{u}} = P_1 + \vec{\boldsymbol{u}} \in \mathcal{F}.$$

Luego también $P_2 \in \mathcal{F} \cup \mathcal{F}_2$.

Nos falta probar que cada $\mathcal{F} \cap \mathcal{F}_i$, para $i = 1, 2$, consta de un sólo punto. Pero esto es consecuencia de que se tiene $\mathcal{F} \cap \mathcal{F}_i \neq \varnothing$ y $\boldsymbol{F} \cap \boldsymbol{F}_i = (\boldsymbol{F}_1 \cap \boldsymbol{F}_2) \cap \boldsymbol{F}_i = \{\vec{\boldsymbol{0}}_E\}$ y del Ejemplo 2.25.

Nótese que si $\vec{\boldsymbol{v}} = \vec{\boldsymbol{v}}_1^* + \vec{\boldsymbol{v}}_2^*$ es otra descomposición de $\vec{\boldsymbol{v}}$ (en caso de que $\boldsymbol{F}_1 \cap \boldsymbol{F}_2 \neq \{\vec{\boldsymbol{0}}_E\}$), se obtendría otra variedad lineal perpendicular común a $\mathcal{F}_1$ y a $\mathcal{F}_2$, así como otros pies:

$$P_1^* = Q_1 + \vec{\boldsymbol{v}}_1^* \in \mathcal{F}_1, \qquad P_2^* = Q_2 - \vec{\boldsymbol{v}}_2^* \in \mathcal{F}_2, \qquad \mathcal{F}^* = P_1^* + \boldsymbol{F}.$$

— Para establecer 4), supongamos que $\mathcal{F}'$ es una variedad lineal, de espacio director $\boldsymbol{F}'$, verificando $\mathcal{F}' \perp \mathcal{F}_i$ y $\mathcal{F}' \cap \mathcal{F}_i = \{P_i'\}$ para $i = 1, 2$ y vamos a demostrar que $\boldsymbol{F}' \subset \boldsymbol{F}$.

$$\mathcal{F}' \perp \mathcal{F}_i \ (i = 1, 2) \Leftrightarrow (\boldsymbol{F}' \subset \boldsymbol{F}_1^\perp \text{ ó } \boldsymbol{F}' \supset \boldsymbol{F}_1^\perp) \text{ y } (\boldsymbol{F}' \subset \boldsymbol{F}_2^\perp \text{ ó } \boldsymbol{F}' \supset \boldsymbol{F}_2^\perp) \Leftrightarrow$$
$$\Leftrightarrow (\boldsymbol{F}' \subset \boldsymbol{F}_1^\perp \text{ y } \boldsymbol{F}' \subset \boldsymbol{F}_2^\perp) \text{ ó } (\boldsymbol{F}' \subset \boldsymbol{F}_1^\perp \text{ y } \boldsymbol{F}' \supset \boldsymbol{F}_2^\perp) \text{ ó } (\boldsymbol{F}' \supset \boldsymbol{F}_1^\perp \text{ y } \boldsymbol{F}' \subset \boldsymbol{F}_2^\perp) \text{ ó}$$
$$(\boldsymbol{F}' \supset \boldsymbol{F}_1^\perp \text{ y } \boldsymbol{F}' \supset \boldsymbol{F}_2^\perp).$$

De estos cuatro posibles casos, sólo el primero es posible, lo cual implica que $\boldsymbol{F}' \subset \boldsymbol{F}_1^\perp \cap \boldsymbol{F}_2^\perp = \boldsymbol{F}$.

Vamos a ver que los otros casos no se pueden dar; ya que si, por ejemplo, si $\boldsymbol{F}' \subset \boldsymbol{F}_1^\perp$ y $\boldsymbol{F}' \supset \boldsymbol{F}_2^\perp$, se tiene que $\dim \boldsymbol{F}' > \dim \boldsymbol{F}_2^\perp = n - \dim \boldsymbol{F}_2$, o sea, $\dim \boldsymbol{F}' + \dim \boldsymbol{F}_2 > n$.

Como hemos supuesto que $\mathcal{F}' \cap \mathcal{F}_2 = \{P_2'\}$, entonces $\dim(\boldsymbol{F}' \cap \boldsymbol{F}_2) = 0$ y, por tanto, $\dim(\boldsymbol{F}' + \boldsymbol{F}_2) = \dim \boldsymbol{F}' + \dim \boldsymbol{F}_2 > n$. Lo cual es una contradicción.

Suponer que uno de los dos restantes casos sea cierto conduciría a una contradicción, siguiendo el mismo razonamiento.

— Para establecer 5), esto es, para ver que la distancia de $\mathcal{F}_1$ a $\mathcal{F}_2$ es $d(P_1, P_2)$, tomemos otros puntos $R_1 \in \mathcal{F}_1$ y $R_2 \in \mathcal{F}_2$.

Como $\overrightarrow{P_1P_2} \perp \overrightarrow{P_1R_1}$, $\overrightarrow{P_1P_2} \perp \overrightarrow{P_2R_2}$ y $\overrightarrow{R_1R_2} = \overrightarrow{R_1P_1} + \overrightarrow{P_2R_2} + \overrightarrow{P_1P_2}$, por el Teorema de Pitágoras se tiene que:

$$\|\overrightarrow{R_1R_2}\|^2 = \|\overrightarrow{R_1P_1} + \overrightarrow{P_2R_2}\|^2 + \|\overrightarrow{P_1P_2}\|^2,$$
$$d(R_1, R_2)^2 = \|\overrightarrow{R_1P_1} + \overrightarrow{P_2R_2}\|^2 + d(P_1, P_2)^2 > d(P_1, P_2)^2.$$

De forma similar de llega (cuando se toman los pies P_1^* y P_2^* considerados arriba) a que $d(P_1^*, P_2^*)$ es la menor distancia entre pares de puntos uno en $\mathcal{F}_1$ y otro en $\mathcal{F}_2$; por lo que $d(P_1, P_2) = d(P_1^*, P_2^*)$.

— Finalmente, supongamos que $\mathcal{F}_1$ y $\mathcal{F}_2$ se cruzan, es decir, $\mathcal{F}_1 \cap \mathcal{F}_2 = \varnothing$ y $\boldsymbol{F}_1 \cap \boldsymbol{F}_2 = \{\vec{\boldsymbol{0}}_E\}$, entonces la variedad lineal perpendicular común $\mathcal{F}$ es única.

En el supuesto que exista otra $\mathcal{F}' = P_1' + F = P_1' + (F_1^\perp + F_2^\perp)$ cuyos pies sean $P_1' \in \mathcal{F}_1$ y $P_2' \in \mathcal{F}_2$, se verifica que $P_1 = P_1'$ y $P_2 = P_2'$; con lo que $\mathcal{F} = \mathcal{F}'$. En efecto, por 5),

$$d(P_1, P_2) = d(P_1', P_2') \Leftrightarrow \|\overrightarrow{P_1P_1'} + \overrightarrow{P_2'P_2}\| = 0 \Leftrightarrow \overrightarrow{P_1P_1'} = \overrightarrow{P_2P_2'}.$$

Como $\overrightarrow{P_1P_1'} = \overrightarrow{P_2P_2'} \in F_1 \cap F_2 = \{\vec{0}_E\}$, se tiene $P_1 = P_1'$ y $P_2 = P_2'$.

Cuando las variedades lineales $\mathcal{F}_1$ y $\mathcal{F}_2$, siendo disjuntas, no se cruzan, existe $\vec{w} \neq \vec{0}$ de $F_1 \cap F_2$. Sean los puntos $Q_1 = P_1 + \vec{w} \in \mathcal{F}_1$ y $Q_2 = P_2 + \vec{w} \in \mathcal{F}_1$, entonces $\overrightarrow{P_2Q_2} = \overrightarrow{P_1Q_1}$ y, por la ley del paralelogramo, $\overrightarrow{P_1P_2} = \overrightarrow{Q_1Q_2}$. Por lo que existe más de una variedades lineales perpendicular común a $\mathcal{F}_1$ y $\mathcal{F}_2$. ◪

6.27 Ejemplo.- *Sean los planos en el espacio euclídeo $\mathbb{R}^4$,*

$$\mathcal{F}_1 = \{(x, y, z, t) \in \mathbb{R}^4 / x + y - 3z + 4 = 0, x + y - 3t + 1 = 0\},$$

$$\mathcal{F}_2 = \{(x, y, z, t) \in \mathbb{R}^4 / t = z, y = 2z\}.$$

Encontrar la variedad perpendicular común a ambos y sus pies.

Consideremos la referencia rectangular canónica en $\mathbb{R}^4$ y adoptaremos las notaciones usadas en la demostración de la Proposición anterior.

Como $\mathcal{F}_1 \cap \mathcal{F}_2 = \varnothing$ (el sistema formado por las ecuaciones $x + y - 3z + 4 = 0, x+y-3t+1 = 0, t = z, y = 2z$ no tiene solución), la suma de sus subespacios directores $F_1 + F_2$ está estrictamente contenido en el espacio vectorial $\mathbb{R}^4$.

Los subespacios vectoriales directores de las variedades lineales $\mathcal{F}_1$ y $\mathcal{F}_2$ y los ortogonales a ellos son:

$$F_1 = \{(x, y, z, t) \in \mathbb{R}^4 / x + y - 3z = 0, x + y - 3t + 1 = 0\} =$$
$$= \{(-u + 3v, u, v, v) \in \mathbb{R}^4 / u, v \in \mathbb{R}\},$$
$$F_2 = \{(x, y, z, t) \in \mathbb{R}^4 / t = z, y = 2z\} = \{(u, 2v, v, v) / u, v \in \mathbb{R}\},$$
$$F_1^\perp = \{(u + v, u + v, -3u, -3v) / u, v \in \mathbb{R}\} = \{(x, y, z, t) / t = -3y - z, x = y\},$$
$$F_2^\perp = \{(0, u + v, -2u, -2v) / u, v \in \mathbb{R}\} = \{(x, y, z, t) / t = -2y - z, x = 0\}.$$

Pues, un vector de componentes (a, b, c, d) pertenece a $F_1^\perp$ si para cualquier vector de F_1 se verifica:

$$(a, b, c, d).(-u + 3v, u, v, v) = (b - a)u + (3a + c + d)v = 0.$$

Relación que se debe verificarse para todo u y v; entonces, $b-a = 0$, $3a+c+d = 0$, cuya soluciones son $a = -(c+d)/3, b = -(c+d)/3$. Análogamente se procede para obtener $F_2^\perp$.

El espacio director de la variedad lineal perpendicular común es:

$$F = (F_1 + F_2)^\perp = F_1^\perp \cap F_2^\perp = \{(0, 0, u, -u) / u \in \mathbb{R}\}.$$

Tomemos puntos en los planos dados:

$$Q_1(-5, 1, 0, -1) \in \mathcal{F}_1, \qquad Q_2(0, 0, 0, 0) \in \mathcal{F}_2,$$

Vamos a descomponer el vector $\overrightarrow{Q_1Q_2} \in \mathbb{R}^4 = (F_1 + F_2) \oplus F$ como suma única de un vector de $F_1 + F_2$ más uno de F.

Dos vectores independientes de F_1 son:

$$\vec{a}_1 = (-1, 1, 0, 0) \qquad \text{y} \qquad \vec{b}_1 = (3, 0, 1, 1).$$

Pues, las componentes de cada uno de ellos satisfacen al sistema $x + y - 3z = 0, x + y - 3t = 0$.

Por otro lado, dos vectores independientes, cuyas componentes satisfacen al sistema $t = z, y = 2z$, de F_2 son:

$$\vec{a}_2 = (1, 0, 0, 0) \qquad \text{y} \qquad \vec{b}_2 = (0, 2, 1, 1).$$

Los vectores $\{\vec{\boldsymbol{a}}_1, \vec{\boldsymbol{a}}_2, \vec{\boldsymbol{b}}_2\}$ forma una base de $\boldsymbol{F}_1 + \boldsymbol{F}_2$ y $\{\vec{\boldsymbol{c}}\} = \{(0,0,1,-1)\}$ es una base de $\boldsymbol{F}$.

$$\overrightarrow{Q_1Q_2} = \lambda\vec{\boldsymbol{a}}_1 + \mu\vec{\boldsymbol{b}}_1 + \nu\vec{\boldsymbol{a}}_2 + \zeta\vec{\boldsymbol{c}},$$

para $\lambda = -1, \mu = 1/2, \nu = 5/2, \xi = -1/2$. Luego,

$$\overrightarrow{Q_1Q_2} = (-\vec{\boldsymbol{a}}_1 + \frac{1}{2}\vec{\boldsymbol{b}}_1) + \frac{5}{2}\vec{\boldsymbol{b}}_2 - \frac{1}{2}\vec{\boldsymbol{c}} = \vec{\boldsymbol{v}}_1 + \vec{\boldsymbol{v}}_2 + \vec{\boldsymbol{u}}.$$

Con $\vec{\boldsymbol{v}}_1 = (5/2, -1, 1/2, 1/2) \in \boldsymbol{F}_1$, $\vec{\boldsymbol{v}}_2 = (5/2, 0, 0, 0) \in \boldsymbol{F}_2$ y en el subespacio ortogonal a $\boldsymbol{F}_1 + \boldsymbol{F}_1$ el vector $\vec{\boldsymbol{u}} = (0, 0, -1/2, 1/2) \in \boldsymbol{F}$.

Los pies (6-1) de la variedad lineal perpendicular común son:

$$P_1 = Q_1 + \vec{\boldsymbol{v}}_1 = \left(-\frac{5}{2}, 0, \frac{1}{2}, -\frac{1}{2}\right), \qquad P_2 = Q_2 - \vec{\boldsymbol{v}}_2 = \left(-\frac{5}{2}, 0, 0, 0\right).$$

Y la perpendicular común es la recta:

$$\mathcal{F} = P_1 + \boldsymbol{F} = \left\{\left(-\frac{5}{2}, 0, u + \frac{1}{2}, -u - \frac{1}{2}\right) \Big/ u \in \mathbb{R}\right\} =$$
$$= \{(x,y,z,t) \in \mathbb{R}^4 / 2x + 5 = 0, y = 0, z + t = 0\}.$$

La distancia entre los planos es:

$$d(\mathcal{F}_1, \mathcal{F}_2) = d(P_1, P_2) = \|\overrightarrow{P_1P_2}\| = \frac{\sqrt{2}}{2}.$$

6.28 Nota.- La distancia entre los planos del Ejemplo anterior se puede determinar sin necesidad de conocer los pies de la perpendicular común, sin más que aplicar la fórmula siguiente [4, Problem 7.13], válida para cualquier par de variedades lineales $\mathcal{F}_1$ y $\mathcal{F}_2$ (de subespacios directores $\boldsymbol{F}_1$ y $\boldsymbol{F}_2$) de un espacio euclídeo de dimensión n y tales que $\mathcal{F}_1 \cap \mathcal{F}_2 = \varnothing$:

$$d(\mathcal{F}_1, \mathcal{F}_2)^2 = \frac{\mathbf{Gram}(\overrightarrow{Q_1Q_2}, \vec{\boldsymbol{a}}_1, \ldots, \vec{\boldsymbol{a}}_m)}{\mathbf{Gram}(\vec{\boldsymbol{a}}_1, \ldots, \vec{\boldsymbol{a}}_m)}. \tag{6-2}$$

Siendo, $Q_1 \in \mathcal{F}_1$, $Q_2 \in \mathcal{F}_2$ y $\{\vec{\boldsymbol{a}}_1, \ldots, \vec{\boldsymbol{a}}_m\}$ una base de $\boldsymbol{F}_1 + \boldsymbol{F}_2$.

Definiendo, además, el gramiano (o determinante de Gram) de los vectores $\vec{\boldsymbol{v}}_1, \ldots, \vec{\boldsymbol{v}}_r$ $(r \leq n)$ como el determinante [(1)]:

$$\mathbf{Gram}(\vec{\boldsymbol{v}}_1, \ldots, \vec{\boldsymbol{v}}_r) = \begin{vmatrix} \|\vec{\boldsymbol{v}}_1\|^2 & \vec{\boldsymbol{v}}_1 \cdot \vec{\boldsymbol{v}}_2 & \cdots & \vec{\boldsymbol{v}}_1 \cdot \vec{\boldsymbol{v}}_r \\ \vec{\boldsymbol{v}}_2 \cdot \vec{\boldsymbol{v}}_1 & \|\vec{\boldsymbol{v}}_2\|^2 & \cdots & \vec{\boldsymbol{v}}_2 \cdot \vec{\boldsymbol{v}}_r \\ \vdots & \vdots & \ddots & \vdots \\ \vec{\boldsymbol{v}}_r \cdot \vec{\boldsymbol{v}}_1 & \vec{\boldsymbol{v}}_r \cdot \vec{\boldsymbol{v}}_2 & \cdots & \|\vec{\boldsymbol{v}}_r\|^2 \end{vmatrix}.$$

(1) Numéricamente, el determinante de Gram coincide con el cuadrado del volumen del paralelepípedo formado por los vectores. En el Ejemplo 5.62, el volumen de tetraedro de vértices los puntos $A(3,2,1), B(1,2,4), C(4,0,3)$ y $D(1,1,7)$ es 1/6 de la raíz cuadrada del gramiano de los vectores $\overrightarrow{AB} = (-2,0,3)$, $\overrightarrow{AC} = (1,-2,2)$, $\overrightarrow{AD} = (-2,-1,6)$,

$$\mathbf{Gram}(\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}) = \begin{vmatrix} 13 & 4 & 22 \\ 4 & 9 & 12 \\ 22 & 12 & 41 \end{vmatrix} = 25.$$

Para el Ejemplo anterior:

$$\mathbf{Gram}(\overrightarrow{Q_1Q_2}, \vec{\boldsymbol{a}}_1, \vec{\boldsymbol{b}}_1, \vec{\boldsymbol{a}}_2) = \begin{vmatrix} 27 & -6 & 16 & 5 \\ -6 & 2 & -3 & -1 \\ 16 & -3 & 11 & 3 \\ 5 & -1 & 3 & 1 \end{vmatrix} = 1.$$

$$\mathbf{Gram}(\vec{\boldsymbol{a}}_1, \vec{\boldsymbol{b}}_1, \vec{\boldsymbol{a}}_2) = \begin{vmatrix} 2 & -3 & -1 \\ -3 & 11 & 3 \\ -1 & 3 & 1 \end{vmatrix} = 2.$$

Para el caso particular de dos recta no coplanarias $\mathfrak{L}$ y $\mathfrak{L}$' con P y Q puntos distintos en $\mathfrak{L}$ y P' y Q' puntos distintos de $\mathfrak{L}$', se tiene que la distancia entre las rectas $\mathfrak{L}$ y $\mathfrak{L}$' es

$$d(\mathfrak{L}, \mathfrak{L}') = \frac{\mathbf{Gram}(\overrightarrow{PP'}, \overrightarrow{PQ}, \overrightarrow{P'Q'})}{\mathbf{Gram}(\overrightarrow{PQ}, \overrightarrow{P'Q'})}. \tag{6-3}$$

6.29 Ejemplo.- *En el espacio euclídeo $\mathbb{R}^3$ y en el caso concreto del Ejemplo que sigue al Ejemplo A.1. podemos ahora determinar los pies de la perpendicular común a los dos rectas dadas (ℓ_1 el eje OZ y la recta ℓ_2 de ecuaciones $x - z - 2 = 0$, $y - z + 1 = 0$) utilizando (6-1) y la distancia entre ellas mediante la fórmula (6-2).*

Utilizando los datos del citado Ejemplo:
$Q_1(0, 0, 0)$, $\vec{\boldsymbol{a}}_1 = (0, 0, 1)$, punto y vector director del eje OZ.
$Q_2(3, 0, 1)$, $\vec{\boldsymbol{a}}_2 = (1, 1, 1)$, punto y vector director de la recta ℓ_2.
Una base de la suma $\boldsymbol{F}_1 + \boldsymbol{F}_2$ de los espacios directores $\boldsymbol{F}_1$ y $\boldsymbol{F}_2$ de las dos rectas es $\{\vec{\boldsymbol{a}}_1, \vec{\boldsymbol{a}}_2\}$. El vector director de la perpendicular común es $\vec{\boldsymbol{c}} = (-1, 1, 0)$. $\overrightarrow{Q_1Q_2}$ se expresa en combinación de $\vec{\boldsymbol{a}}_1$, $\vec{\boldsymbol{a}}_2$ y $\vec{\boldsymbol{c}}$ por:

$$\overrightarrow{Q_1Q_2} = -\frac{1}{2}\vec{\boldsymbol{a}}_1 + \frac{3}{2}\vec{\boldsymbol{a}}_2 - \frac{3}{2}\vec{\boldsymbol{c}}.$$

Poniendo $\vec{\boldsymbol{v}}_1 = -\frac{1}{2}\vec{\boldsymbol{a}}_1 \in \boldsymbol{F}_1$ y $\vec{\boldsymbol{v}}_2 = \frac{3}{2}\vec{\boldsymbol{a}}_2 \in \boldsymbol{F}_2$, por (6-1), los pies de la perpendicular común son:

$$P_1 = Q_1 + \vec{\boldsymbol{v}}_1 = \left(0, 0, -\frac{1}{2}\right), \qquad P_2 = Q_2 - \vec{\boldsymbol{v}}_2 = \left(\frac{3}{2}, -\frac{3}{2}, -\frac{1}{2}\right).$$

La mínima distancia ente las rectas dadas es $\|\overrightarrow{P_1P_2}\| = 3/\sqrt{2}$. O bien, utilizando la fórmula (6-2), sin necesidad de haber obtenido los pies de la recta perpendicular común:

$$d(\ell_1, \ell_2)^2 = \frac{\mathbf{Gram}(\overrightarrow{Q_1Q_2}, \vec{\boldsymbol{a}}_1, \vec{\boldsymbol{a}}_2)}{\mathbf{Gram}(\vec{\boldsymbol{a}}_1, \vec{\boldsymbol{a}}_2)} = \frac{\begin{vmatrix} 10 & 1 & 4 \\ 1 & 1 & 1 \\ 4 & 1 & 3 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix}} = \frac{9}{2}.$$

Los ejemplos que hemos expuesto, relativos a cálculos de distancias, se pueden enfocar de una forma más directa en el caso del espacio euclídeo canónico $\mathbb{R}^n$; no teniendo que determinar previamente vectores independientes en

el subespacio director, para el caso de distancia de un punto a una variedad lineal, como se pone de manifiesto en el siguiente resultado (2):

6.30 Proposición.- *Dado un punto $P \in \mathbb{R}^n$ y una variedad lineal $\mathcal{F}$ en $\mathbb{R}^n$ dada por un sistema de ecuaciones lineales*

$$\begin{array}{rcl} a_1^1x^1 + a_2^1x^2 + \cdots + a_n^1x^n & = & b^1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & \cdots \\ a_1^mx^1 + a_2^mx^2 + \cdots + a_n^mx^n & = & b^m \end{array}$$

o en forma matricial $AX = B$, donde

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ \cdots & \cdots & \cdots & \cdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix}, \ B = \begin{pmatrix} b^1 \\ \vdots \\ b^m \end{pmatrix}, \ X = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$$

Se supone que $m \leq n$ y que $\operatorname{rango}(A) = m$, i.e. el sistema de ecuaciones es compatible indeterminado y define la variedad lineal $\mathcal{F}$ de dimensión $n - m$ en $\mathbb{R}^n$.

Consideremos la matriz cuadrada de orden $m + 1$:

$$M = \begin{pmatrix} A\ {}^tA & AP - B \\ {}^t(AP - B) & 0 \end{pmatrix}.$$

Entonces, la distancia de P a la variedad lineal $\mathcal{F}$ es:

$$d(P, \mathcal{F}) = \sqrt{-\frac{\det M}{\det(A\ {}^tA)}}. \tag{6-4}$$

Nótese que el radicando no es negativo debido a la condición impuesta sobre el $\operatorname{rango}(A)$ y las propiedades del gramiano $\det(A\ {}^tA) > 0$.

6.31 Ejemplo.- *Distancia del punto $P(1, 1, 1, 1)$ a la variedad lineal de $\mathbb{R}^4$: $3x + y - z + t = 1, \quad x - 2y + z + 2t = 2.$*

Como aplicación de los métodos expuestos para hallar distancias entre variedades lineales, vamos a encontrar la solución, $3\sqrt{5/58}$, utilizando los tres dados.

— Primero, determinando los pies de la variedad perpendicular común.

Uno de ellos es P, de la variedad lineal $\mathcal{F}_1 = \{P\}$. En este caso, debemos determinar la proyección ortogonal de P sobre $\mathcal{F}_2$. El subespacio director $\boldsymbol{F}$ de la variedad ortogonal a la dada (que denotamos por $\mathcal{F}_2$) por P está formado por los vectores ortogonales a los dos vectores generadores del subespacio director $\boldsymbol{F}_2$ de $\mathcal{F}_2$.

Tomamos los vectores $\vec{\boldsymbol{a}}_2$ y $\vec{\boldsymbol{b}}_2$ cuyas componentes se dan más abajo, en el segundo método de resolución; las componentes $(\xi^1, \xi^2, \xi^3, \xi^4)$ de un vector perpendicular a $\vec{\boldsymbol{a}}_2$ y a $\vec{\boldsymbol{b}}_2$, deben satisfacer:

$$\xi^1 + \frac{5}{3}\xi^3 - \frac{4}{3}\xi^4 = 0, \quad \xi^2 + \frac{4}{3}\xi^3 + \frac{1}{3}\xi^4 = 0.$$

(2) Alexei Uteshev.- Distance evaluation between geometric objects http://pmpu.ru/vf4/matricese/optimize/distancee (16-8-2011)

Resolviendo, con variables indeterminadas ξ^1 y ξ^2, se obtiene que las coordenadas de los puntos de la variedad lineal perpendicular a $\mathfrak{F}_2$ por el punto $P(1,1,1,1)$ son:

$$(1,1,1,1)+\left(\xi^1,\xi^2,-\frac{\xi^1}{7}-\frac{4\xi^2}{7},\frac{4\xi^1}{7}-\frac{5\xi^2}{7}\right).$$

Sustituyendo en las ecuaciones de $\mathfrak{F}_2$ y resolviendo en las variables ξ^1 y ξ^2, se obtiene el pie Q_2 de dicha variedad perpendicular y la distancia de P a $\mathfrak{F}$ son:

$$P_2\left(\frac{8}{29},\frac{37}{58},\frac{38}{29},\frac{49}{58}\right),\qquad d(P,\mathfrak{F}_2)=\|\overrightarrow{PQ_2}\|=3\sqrt{\frac{5}{58}}.$$

— <u>Segundo</u>, por la fórmula (6-2), en términos de gramianos.

Necesitamos encontrar un punto en la variedad lineal dada, que hemos denotamos por $\mathfrak{F}_2$, y dos vectores independientes de su subespacio director $\boldsymbol{F}_2$.

La solución general del sistema que determina sus puntos es:

$$z=\frac{5}{3}x+\frac{4}{3}y,\qquad t=1-\frac{4}{3}x+\frac{1}{3}y.$$

Con lo que la podemos parametrizar de la forma:

$$(x,y,z,t)=(0,0,0,1)+\lambda\Big(1,0,\frac{5}{3},-\frac{4}{3}\Big)+\mu\Big(0,1,\frac{4}{3},\frac{1}{3}\Big).$$

Así, tenemos $Q_2(0,0,0,1)\in\mathfrak{F}_2$ y los vectores independientes de su espacio director$^{(*)}$:

$$\vec{\boldsymbol{a}}_2=\Big(1,0,\frac{5}{3},-\frac{4}{3}\Big),\qquad \vec{\boldsymbol{b}}_2=\Big(0,1,\frac{4}{3},\frac{1}{3}\Big).$$

Y obtenemos:

$$d(P,\mathfrak{F}_2)^2=\frac{\mathbf{Gram}(\overrightarrow{PQ_2},\vec{\boldsymbol{a}}_2,\vec{\boldsymbol{b}}_2)}{\mathbf{Gram}(\vec{\boldsymbol{a}}_2,\vec{\boldsymbol{b}}_2)}=\frac{\begin{vmatrix}3&\frac{8}{3}&\frac{7}{3}\\\frac{8}{3}&\frac{50}{9}&\frac{16}{9}\\\frac{7}{3}&\frac{16}{9}&\frac{26}{9}\end{vmatrix}}{\begin{vmatrix}\frac{50}{9}&\frac{16}{9}\\\frac{16}{9}&\frac{26}{9}\end{vmatrix}}=\frac{45}{58}.$$

— Y <u>Tercero</u>, el más rápido y directo, utilizando la fórmula (6-4).

$$A=\begin{pmatrix}3&1&-1&1\\1&-2&1&2\end{pmatrix},\ B=\begin{pmatrix}1\\2\end{pmatrix},\ P=\begin{pmatrix}1\\1\\1\\1\end{pmatrix},\ AP=\begin{pmatrix}4\\2\end{pmatrix},$$

$$AP-B=\begin{pmatrix}3\\0\end{pmatrix},\ A\,{}^{\mathrm{t}}\!A=\begin{pmatrix}12&2\\2&10\end{pmatrix},\quad M=\begin{pmatrix}12&2&3\\2&10&0\\3&0&0\end{pmatrix}.$$

$$d(P,\mathfrak{F}_2)=\sqrt{-\frac{\det M}{\det(A\,{}^{\mathrm{t}}\!A)}}=3\sqrt{\frac{5}{58}}.$$

(*)

Mathematica proporciona dos comandos, `LinearSolve` y `NullSpace`, que nos permiten determinar un punto de una variedad lineal, definida por un sistema lineal $AX = B$, y una base de vectores para su subespacio director. En este ejemplo concreto,

`LinearSolve[{{3,1,-1,1}, {1,-2,1,2}},{1,2}]`

nos da el punto $(4/7, -5/7, 0, 0)$ de $\mathcal{F}_2$ (si el sistema no es compatible, el programa nos lo notifica). La instrucción:

`NullSpace[{{3,1,-1,1}, {1,-2,1,2}}]`

da una base $\{(-4, 5, 0, 7), (1, 4, 7, 0)\}$ de F_2.

La fórmula (6-2), usando estos datos, nos da:

$$d(P, \mathcal{F}_2)^2 = \frac{\begin{vmatrix} \frac{251}{49} & -\frac{97}{7} & -\frac{100}{7} \\ -\frac{97}{7} & 90 & 16 \\ -\frac{100}{7} & 16 & 66 \end{vmatrix}}{\begin{vmatrix} 90 & 16 \\ 16 & 66 \end{vmatrix}} = \frac{45}{58}.$$

Las rutinas personales siguientes nos valen para calcular la distancia de un punto a una variedad lineal :

```
gram[lista_List]:=Module[{long},
long=Table[{Length[lista[[i]]]},{i,1,Length[lista]}];
If[Length[Union[long]]!=1,
"Los vectores son de distinta longitud",
Table[Map[(lista[[i]].#)&,lista],{i,1,Length[lista]}]]];


distanciaPVariedadLineal[ P_List,  {A_List, B_List} ]:=
Module[{solSA,solSH},
solSA=Flatten[LinearSolve[A,B],1];
solSH=NullSpace[A];
Sqrt[Det[gram[Append[solSH,P-solSA]]]/ Det[gram[solSH]]]];
```

(Estas rutinas, y otra que determina la distancia entre variedades lineales de $\mathbb{R}^n$, se pueden descargar de

http://webpages.ull.es/users/amontes/math/dvlrn.nb

que contiene ejemplos.)

Una vez cargadas ambas, en el ejemplo en cuestión conocido P y las ecuaciones cartesianas $AX = B$ de la variedad lineal, basta sustituir en la rutina los valores de P, A y B, poniendo `distanciaPVariedadLineal[P, {A, B}]`, es decir:

`distanciaPVariedadLineal[{1,1,1,1},{{{3,1,-1,1}, {1,-2,1,2}},{1,2}}]`

para obtener: $3\sqrt{\frac{5}{58}}$.

6.32 Ejercicio.- Obtener que la distancia entre las variedades lineales de $\mathbb{R}^5$ cuyas ecuaciones paramétricas se dan es $d = 150$,

$$(x^1, x^2, x^3, x^4, x^5) = (89, 37, 111, 13, 54) + \lambda_1(1, 1, 0, -1, -1) + \mu_1(1, -1, 0, -1, 1),$$

$$(x^1, x^2, x^3, x^4, x^5) = (42, -16, -39, 71, 3) + \lambda_2(1, 1, 0, 1, 1) + \mu_2(1, -1, 0, 1, -1).$$

6.33 Ejemplo.- *La distancia entre hiperplanos paralelos con ecuaciones* $\mathcal{H}_1 : a_1x^1 + \cdots + a_nx^n = b_1, \quad \mathcal{H}_2 : a_1x^1 + \cdots + a_nx^n = b_2,$

respecto a una referencia rectangular $\{O; \vec{\boldsymbol{e}}_1, \dots \vec{\boldsymbol{e}}_n\}$, *es*

$$d(\mathcal{H}_1, \mathcal{H}_2) = \frac{|b_1 - b_2|}{\sqrt{a_1^2 + \cdots + a_n^2}}.$$

Como $\vec{\boldsymbol{a}} = (a_1, \dots, a_n)$ son las componentes de un vector perpendicular a ambos hiperplanos, la recta $\mathcal{L} = O + \widetilde{\{\vec{\boldsymbol{a}}\}}$, interseca a cada hiperplano en un solo punto, sean $P_1 = O + \lambda_1 \vec{\boldsymbol{a}} \in \mathcal{H}_1$, $P_2 = O + \lambda_2 \vec{\boldsymbol{a}} \in \mathcal{H}_2$. Así:

$$\overrightarrow{P_1P_2} \in \mathcal{L} = \mathcal{H}_1^{\perp} = \mathcal{H}_2^{\perp} \quad \text{y} \quad d(\mathcal{H}_1, \mathcal{H}_2) = d(P_1, P_2).$$

Sustituyendo las coordenadas de los puntos P_1 y P_2 en los respectivos hiperplanos que los contienen, resulta $\lambda_1 \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}} = b_1$ y $\lambda_2 \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}} = b_2$. Luego,

$$\overrightarrow{P_1P_2} = \overrightarrow{P_1O} + \overrightarrow{OP_2} = \frac{b_2 - b_1}{\|\vec{\boldsymbol{a}}\|^2} \vec{\boldsymbol{a}} \qquad \Rightarrow \qquad d(\mathcal{H}_1, \mathcal{H}_2) = \frac{|b1 - b_2|}{\|\vec{\boldsymbol{a}}\|}.$$

6.3 Isometrías

Estudiamos a continuación las aplicaciones entre espacios euclídeos que conservan las distancias; ellas son necesariamente aplicaciones afines, con aplicaciones lineales asociadas ortogonales (inyectivas). Esto significa que dichas aplicaciones son compatibles tanto con las estructuras afines como con las euclídeas de los espacios entre las que están definidas.

Existen diferentes denominaciones a la hora de nombrar a estas aplicaciones en general, también cuando son biyectivas (lo que ocurre si los espacios son de la misma dimensión, finita) y cuando el espacio de partida y el de llegada coinciden. Nosotros adoptaremos, para nombrarlas, las que nos parecen que figuran con más asiduidad en la literatura al respecto.

Definición y caracterizaciones

6.34 Definición.- *Se llama* ***aplicación isométrica*** *a una aplicación afín* $f\colon \mathcal{A} \to \mathcal{A}'$ *entre espacios euclídeos que conserva las distancias, es decir, tal que para todo* $P, Q \in \mathcal{A}$

$$d(f(P), f(Q)) = d(P, Q).$$

Cuando $\mathcal{A}$ *y* $\mathcal{A}'$ *son de la misma dimensión (finita), la aplicación isométrica se dice que es una* ***isometría***. *Si los espacio coinciden se le da el nombre de* ***movimiento rígido*** *(o simplemente,* ***movimiento****).*

Se ha denotado la función distancia en los dos espacios euclídeos con la misma letra d, sin embargo, ello no creará confusión si se tiene en cuenta en qué espacio están los puntos a los que se les está aplicando la función distancia.

El siguiente resultado da una caracterización para las aplicaciones afines que son isométricas.

6.35 Proposición.- *Sean* $(\mathcal{A}, E)$ *y* $(\mathcal{A}', E')$ *espacios euclídeos y* $f\colon \mathcal{A} \to \mathcal{A}'$ *una aplicación afín de aplicación lineal asociada* $\tilde{f}$, *entonces los siguientes apartados son equivalentes:*

1) *f es una aplicación isométrica (es decir, f conserva la distancia).*
2) *$\forall \vec{\boldsymbol{v}} \in E$, $\|\tilde{f}(\vec{\boldsymbol{v}})\| = \|\vec{\boldsymbol{v}}\|$ (es decir, $\tilde{f}$ conserva el módulo de vectores).*
3) *$\forall \vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \in E$, $\tilde{f}(\vec{\boldsymbol{u}}) \cdot \tilde{f}(\vec{\boldsymbol{v}}) = \vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}}$ (es decir, $\tilde{f}$ es una isometría lineal, conserva el producto escalar).*

Demostración.-
$(1 \Rightarrow 2)$ Para un vector $\vec{\boldsymbol{v}} \in E$, sean $P, Q \in \mathcal{A}$ tales que $\vec{\boldsymbol{v}} = \overrightarrow{PQ}$, entonces
$$\|\tilde{f}(\vec{\boldsymbol{v}})\| = \|\tilde{f}(\overrightarrow{PQ})\| = \|\overrightarrow{f(P)f(Q)}\| = d(f(P), f(Q)) = d(P, Q) = \|\vec{\boldsymbol{v}}\|.$$

$(2 \Rightarrow 3)$ Para $\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \in E$, $\|\vec{\boldsymbol{u}} - \vec{\boldsymbol{v}}\|^2 = \vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{u}} + \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{v}} - 2\vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}}$.
Por otro lado, como $\tilde{f}$ conserva los módulos de los vectores y es lineal:
$$\|\vec{\boldsymbol{u}} - \vec{\boldsymbol{v}}\|^2 = \|\tilde{f}(\vec{\boldsymbol{u}} - \vec{\boldsymbol{v}})\|^2 = \|\tilde{f}(\vec{\boldsymbol{u}}) - \tilde{f}(\vec{\boldsymbol{v}})\|^2 =$$
$$= \tilde{f}(\vec{\boldsymbol{u}}) \cdot \tilde{f}(\vec{\boldsymbol{u}}) + \tilde{f}(\vec{\boldsymbol{v}}) \cdot \tilde{f}(\vec{\boldsymbol{v}}) - 2\tilde{f}(\vec{\boldsymbol{u}}) \cdot \tilde{f}(\vec{\boldsymbol{v}}) == \vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{u}} + \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{v}} - 2\tilde{f}(\vec{\boldsymbol{u}}) \cdot \tilde{f}(\vec{\boldsymbol{v}}).$$
Comparando esta dos últimas relaciones obtenidas, se tiene que
$$\vec{\boldsymbol{u}} \cdot \vec{\boldsymbol{v}} = \tilde{f}(\vec{\boldsymbol{u}}) \cdot \tilde{f}(\vec{\boldsymbol{v}}).$$

$(3 \Rightarrow 1)$ Sean $P, Q \in \mathcal{A}$,
$$d\big(f(P), f(Q)\big)^2 = \|\overrightarrow{f(P)f(Q)}\|^2 = \|\tilde{f}(\overrightarrow{PQ})\|^2 =$$
$$= \tilde{f}(\overrightarrow{PQ}) \cdot \tilde{f}(\overrightarrow{PQ}) = \overrightarrow{PQ} \cdot \overrightarrow{PQ} = \|\overrightarrow{PQ}\|^2 = d(P, Q)^2.$$
◪

Como consecuencia de estas caracterizaciones de las aplicaciones afines que son movimientos y de que las isometrías lineales son biyectivas (Definición 5.69), se tiene que:

6.36 Proposición.- *Los movimientos son afinidades.* ◪

En la definición de aplicación isométrica que hemos dado exigimos, de antemano, que la aplicación sea afín; no obstante, este requisito no es necesario, pues una aplicación $f : \mathcal{A} \to \mathcal{A}$' que conserva la distancia es de hecho una aplicación afín, como establecemos a continuación.

6.37 Proposición.- *Sean $(\mathcal{A}, E)$ y $(\mathcal{A}', E')$ espacios euclídeos y $f : \mathcal{A} \to \mathcal{A}$ una aplicación $\Rightarrow$ $\big(f$ conserva la distancia $\Leftrightarrow$ f es una aplicación afín y su aplicación lineal asociada conserva el producto escalar$\big)$.*

Demostración.- $(\Rightarrow)$ Supongamos que f conserva la distancia; fijado un punto $P \in \mathcal{A}$, para cada vector $\vec{\boldsymbol{u}} \in E$, existe un único $X \in \mathcal{A}$ tal que $\vec{\boldsymbol{u}} = \overrightarrow{PX}$. Definimos la aplicación
$$\hat{f} : E \to E, \qquad \vec{\boldsymbol{u}} = \overrightarrow{PQ} \mapsto \hat{f}(\vec{\boldsymbol{u}}) = \overrightarrow{f(P)f(X)}.$$
Probemos que $\hat{f}$ conserva el producto escalar. Tomemos dos vectores $\vec{\boldsymbol{u}} = \overrightarrow{PX}$ y $\vec{\boldsymbol{v}} = \overrightarrow{PY}$ y expresemos separadamente $d(X, Y)$ y $d(f(X), f(Y))$ en términos de $\vec{\boldsymbol{u}}$ y $\vec{\boldsymbol{v}}$.
$$d(X, Y)^2 = \|\overrightarrow{XY}\|^2 = \|\overrightarrow{XP} + \overrightarrow{PY}\|^2 = \|\overrightarrow{XP}\|^2 + \|\overrightarrow{PY}\|^2 + 2\overrightarrow{XP} \cdot \overrightarrow{PY} =$$

$$= d(X,P)^2 + d(P,Y)^2 - 2\vec{\boldsymbol{u}}\cdot\vec{\boldsymbol{v}}.$$

Así mismo,

$d(f(X),f(Y))^2 = d(f(X),f(P))^2 + d(f(P),f(Y))^2 - 2\overrightarrow{f(P)f(X)}\cdot\overrightarrow{f(P)f(Y)}.$

De la igualdad de los primeros miembros, se sigue que:

$$\vec{\boldsymbol{u}}\cdot\vec{\boldsymbol{v}} = \overrightarrow{f(P)f(X)}\cdot\overrightarrow{f(P)f(Y)} = \hat{f}(\vec{\boldsymbol{u}})\cdot\hat{f}(\vec{\boldsymbol{v}}).$$

Es decir, $\hat{f}$ conserva el producto escalar y, por tanto (pág.113), es lineal.

Nos falta ver que es la aplicación lineal asociada a f. En efecto,

$\hat{f}(\overrightarrow{XY}) = \hat{f}(\overrightarrow{PY}-\overrightarrow{PX}) = \hat{f}(\overrightarrow{PY}) - \hat{f}(\overrightarrow{PX}) = \overrightarrow{f(P)f(Y)} - \overrightarrow{f(P)f(X)} = \overrightarrow{f(X)f(Y)}.$

($\Leftarrow$) Sea f una aplicación afín cuya aplicación lineal asociada $\tilde{f}$ conserva el producto escalar, entonces, para puntos $X,Y\in\mathcal{A}$:

$$d\big(f(X),f(Y)\big) = \|\overrightarrow{f(X)f(Y)}\| = \|\tilde{f}(\overrightarrow{XY})\| = \|\overrightarrow{XY}\| = d(X,Y).$$

Se concluye que f es una aplicación isométrica. ◾

6.38 Ejemplo.- *Las traslaciones y las simetrías ortogonales son movimientos.*

— Una traslación es una aplicación afín con aplicación lineal asociada la identidad, que obviamente es una isometría lineal. O bien, directamente a partir de la definición, $\|\overrightarrow{t_{\vec{\boldsymbol{v}}}(P)t_{\vec{\boldsymbol{v}}}(Q)}\| = \|\overrightarrow{(P+\vec{\boldsymbol{v}})(Q+\vec{\boldsymbol{v}})}\| = \|\overrightarrow{PQ}\|$.

— Si $s_{\mathcal{F}}$ es la simetría ortogonal con respecto a $\mathcal{F}$, se tiene que su aplicación lineal asociada es la simetría s_F respecto a F en la dirección de $F^\perp$. Entonces, para $\vec{\boldsymbol{v}}\in E$:

$$\vec{\boldsymbol{v}} = p_F(\vec{\boldsymbol{v}}) + p_{F^\perp}(\vec{\boldsymbol{v}}), \qquad s_F(\vec{\boldsymbol{v}}) = p_F(\vec{\boldsymbol{v}}) - p_{F^\perp}(\vec{\boldsymbol{v}}).$$

Como $p_F(\vec{\boldsymbol{v}})\cdot p_{F^\perp}(\vec{\boldsymbol{v}}) = 0$ y por la propiedad 6) de la norma (Proposición 5.21), se tiene que

$$\|s_F(\vec{\boldsymbol{v}})\| = \|(\vec{\boldsymbol{v}})\|.$$

Es decir, s_F conserva la norma de vectores y, por tanto, $s_{\mathcal{F}}$ es un movimiento.

Con el objetivo de profundizar en el estudio de las isometrías, es importante investigar los puntos fijos de las aplicaciones afines.

En la siguiente proposición se establece que todo movimiento es la composición (conmutativa) de una traslación con un movimiento con algún punto fijo. Este resultado puede ser visto como una generalización del teorema de Chasles (3) sobre los movimientos rígidos en $\mathbb{R}^3$.

6.39 Proposición.- *Sea una espacio euclídeos $(\mathcal{A},E)$ de dimension n. Para todo movimiento $f\colon\mathcal{A}\to\mathcal{A}$ existe una único movimiento $g\colon\mathcal{A}\to\mathcal{A}$ y una única traslación $t_{\vec{\boldsymbol{w}}}$, con $\tilde{f}(\vec{\boldsymbol{w}}) = \vec{\boldsymbol{w}}$, tal que los puntos fijos de g, $\mathcal{G} = \{X\in\mathcal{A}/g(X) = X\}$, es una variedad lineal no vacía de $\mathcal{A}$ de subespacio director $G = \mathrm{Ker}(\tilde{f} - 1_E)$ y verificándose que*

$$f = t_{\vec{\boldsymbol{w}}}\circ g = g\circ t_{\vec{\boldsymbol{w}}}.$$

(3) Teorema de Chasles: Un desplazamiento espacial de un cuerpo rígido puede ser definido por una rotación alrededor de una recta y una traslación a lo largo de la misma recta, llamado "desplazamiento del tornillo" o "movimiento helicoidal".

Además, se verifican las siguientes propiedades adicionales:

1) $f = g$ *y* $\vec{w} = \vec{0} \Leftrightarrow \mathcal{F} = \{X \in \mathcal{A} / f(X) = X\} \neq \varnothing$.

2) $\mathcal{F} = \{X \in \mathcal{A} / f(X) = X\} = \varnothing \Rightarrow \dim \boldsymbol{G} \geq 1$.

Nótese que la condición $\tilde{f}(\vec{w}) = \vec{w}$ nos dice que $\vec{w} \in \operatorname{Ker}(\tilde{f} - 1_E)$ (subespacio de vectores propios correspondiente al valor propio $1_{\mathbb{R}}$ de $\tilde{f}$) y también que, para cualquier $X \in \mathcal{A}$, $\overrightarrow{f(X)f(X+\vec{w})} = \vec{w}$. Además, la dirección de la traslación $t_{\vec{w}}$ es paralela a la variedad lineal de los puntos fijos del movimiento g.

Demostración.- El vector $\vec{w}$, que define la traslación única a buscar, está determinado por un punto $P \in \mathcal{A}$ y su imagen $f(P)$, es decir $\vec{w} = \overrightarrow{Pf(P)}$. Para hallar el punto P, tomemos un punto $Q \in \mathcal{A}$ que no sea fijo para f ($f(Q) \neq Q$); como, por la Proposición 5.76,

$$\boldsymbol{E} = \operatorname{Ker}(\tilde{f} - 1_E) \oplus \operatorname{Im}(\tilde{f} - 1_E),$$

el vector $\overrightarrow{Qf(Q)}$ se descompone en forma única:

$$\overrightarrow{Qf(Q)} = \vec{w} + (\tilde{f}(\vec{v}) - \vec{v}),$$

donde $\vec{w} \in \operatorname{Ker}(\tilde{f} - 1_E)$ es único y $\vec{v} \in \mathcal{A}$.

El vector $\vec{w}$ no depende del punto Q tomado, pues para cualquier otro punto $X \in \mathcal{A}$:

$$\overrightarrow{Xf(X)} = \overrightarrow{XQ} + \overrightarrow{Qf(Q)} + \overrightarrow{f(Q)f(X)} =$$

$$= \overrightarrow{XQ} + \vec{w} + \tilde{f}(\vec{v}) - \vec{v} + \tilde{f}(\overrightarrow{QX}) = \vec{w} + (\tilde{f} - 1_E)(\vec{v} + \overrightarrow{QX}).$$

Luego, $\overrightarrow{Xf(X)}$ se descompone en forma única en la suma del vector $\vec{w} \in \operatorname{Ker}(\tilde{f} - 1_E)$ más un vector de $\operatorname{Im}(\tilde{f} - 1_E)$, y esto se verifica para todo $X \in \mathcal{A}$ [(4)].

Tomando $P \in \mathcal{A}$ tal que $\overrightarrow{QP} = -\vec{v}$ (es decir, $P = Q - \vec{v}$) se tiene

$$\vec{w} = \overrightarrow{Pf(P)}.$$

Obsérvese que aunque $\vec{w}$ es único, el punto P encontrado no es único; pues puede tomarse otro $\vec{v}'$ tal que $\tilde{f}(\vec{v}) - \vec{v} = \tilde{f}(\vec{v}') - \vec{v}'$.

Determinado el vector $\vec{w}$, definimos ahora $g: \mathcal{A} \to \mathcal{A}$ por:

$$g = t_{-\vec{w}} \circ f.$$

Se verifican todas las afirmaciones del enunciado:

— Como composición de movimientos, g es un movimiento.

— Como $\vec{w} \in \operatorname{Ker}(\tilde{f} - 1_E)$, $\tilde{f}(\vec{w}) = \vec{w}$.

---

(4) Es importante hacer notar que si f tiene algún punto fijo, sea R, entonces

$$\vec{0}_E = \overrightarrow{Rf(R)} = \vec{w} + (\tilde{f} - 1_E)(\vec{v} + \overrightarrow{QR}),$$

y, debido a que $\operatorname{Ker}(\tilde{f} - 1_E) \oplus \operatorname{Im}(\tilde{f} - 1_E) = \boldsymbol{E}$, esta descomposición es única, con lo que $\vec{w} = \vec{0}$. Esto lo utilizaremos para establecer la afirmación 1) del enunciado.

— Las aplicaciones lineales asociadas de f y g coinciden, es decir, $\tilde{f} = \tilde{g}$. En efecto, para todo $X \in \mathcal{A}$ se tiene:
$g(X) = g(P) + \tilde{g}(\overrightarrow{PX}) = P + \tilde{g}(\overrightarrow{PX}), f(X) - \vec{\boldsymbol{w}} = f(P) + \tilde{f}(\overrightarrow{PX}) - \vec{\boldsymbol{w}} = P + \tilde{f}(\overrightarrow{PX})$.
Al ser los primeros miembros de estas relaciones iguales, también lo son los últimos miembros de ellas. Por lo que, $\tilde{f} = \tilde{g}$.

— Como $g(P) = f(P) - \vec{\boldsymbol{w}} = f(P) + \overrightarrow{f(P)P} = P$, ocurre que el conjunto de puntos fijos de g, $\mathcal{G} = \{X \in \mathcal{A}/g(X) = X\} \neq \varnothing$ y por la Proposición 4.7, $\mathcal{G}$ es una variedad lineal con subespacio director $\mathrm{Ker}(\tilde{g} - 1_E) = \mathrm{Ker}(\tilde{f} - 1_E)$ (subespacio vectorial de vectores propio correspondiente al valor propio $1_{\mathbb{R}}$ de $\tilde{f}$).

— Claramente $f = t_{\vec{\boldsymbol{w}}} \circ g$.

— $t_{\vec{\boldsymbol{w}}} \circ g = g \circ t_{\vec{\boldsymbol{w}}}$; en efecto, para cualquier $X \in \mathcal{A}$:
$(g \circ t_{\vec{\boldsymbol{w}}})(X) = g(X + \vec{\boldsymbol{w}}) = g(X) + \tilde{g}(\vec{\boldsymbol{w}}) = f(X) - \vec{\boldsymbol{w}} + \tilde{f}(\vec{\boldsymbol{w}}) = f(X)$.

— Unicidad de la descomposición.

El procedimiento que hemos seguido nos ha permitido obtener sin ambigüedad un vector $\vec{\boldsymbol{w}}$ y un movimiento que nos permite descomponer $f = t_{\vec{\boldsymbol{w}}} \circ g$, verificándose las propiedades requeridas. En el supuesto que existiera otra forma de obtener tal descomposición de f, $f = t_{\vec{\boldsymbol{w}}'} \circ g'$, se debe verificar que $\vec{\boldsymbol{w}}' = \vec{\boldsymbol{w}}$ y $g = g'$. Basta con demostrar que los vectores coinciden.

Usaremos que f y g' tiene la misma aplicación lineal asociada, $\tilde{f} = \tilde{g}'$, que $\vec{\boldsymbol{w}}' \in \mathrm{Ker}(\tilde{f} - 1_E)$ y que g' tiene puntos fijos; sea S un punto fijo de g'.

$$f(S) = g'(S) + \vec{\boldsymbol{w}}' = S + \vec{\boldsymbol{w}}' \Rightarrow \overrightarrow{Sf(S)} = \vec{\boldsymbol{w}}' \in \mathrm{Ker}(\tilde{f} - 1_E).$$

$$\vec{\boldsymbol{w}} - \vec{\boldsymbol{w}}' = \overrightarrow{Pf(P)} - \overrightarrow{Sf(S)} = \overrightarrow{PS} + \overrightarrow{Sf(P)} - \overrightarrow{Sf(P)} - \overrightarrow{f(P)f(S)} =$$

$$= \overrightarrow{PS} - \tilde{f}(\overrightarrow{PS}) = -(f - 1_E)(\overrightarrow{PS}) \in \mathrm{Im}(f - 1_E).$$

$$\vec{\boldsymbol{w}} - \vec{\boldsymbol{w}}' \in \mathrm{Ker}(f - 1_E) \cap \mathrm{Im}(f - 1_E) = \{0_E\}.$$

Propiedades adicionales:

— 1) $f = g$ y $\vec{\boldsymbol{w}} = \vec{\boldsymbol{o}} \Leftrightarrow \mathcal{F} = \{X \in \mathcal{A}/f(X) = X\} \neq \varnothing$.
Si $f = g$ y $\vec{\boldsymbol{w}} = \vec{\boldsymbol{o}}$, la descomposición única de f es $f = t_{\vec{\boldsymbol{o}}} \circ g = g$; además, el conjunto de puntos fijos de $g(= f)$ es no vacío.
La demostración de la otra implicación figura en la nota al pie de la página 140.

— 2) $\mathcal{F} = \{X \in \mathcal{A}/f(X) = X\} = \varnothing \Rightarrow \dim \boldsymbol{G} \geq 1$.
Es equivalente a demostrar: $\dim \boldsymbol{G} < 1 \Rightarrow \mathcal{F} = \{X \in \mathcal{A}/f(X) = X\} \neq \varnothing$. Pero esto está establecido en la Proposición 4.7: f tienen único punto fijo $\Leftrightarrow \mathrm{Ker}(\tilde{f} - 1_E) = \{\vec{\boldsymbol{0}}_E\}$. ■

6.4 Movimientos en espacios euclídeos de dimensiones dos y tres

En el plano euclidiano afín los movimientos (además de la identidad) se clasifican de la siguiente manera. Un movimiento f que tiene algún punto fijo es una rotación si es un movimiento directo, de lo contrario, es una simetría respecto una recta. Si f no tiene puntos fijos, entonces o bien es una traslación no trivial o la composición de una simetría respecto una recta con una traslación no trivial paralelamente a esta recta.

En un espacio euclídeo de dimensión 3, los movimientos (además de la identidad) se clasifican de la siguiente manera. Hay tres tipos de movimientos que tienen algún punto fijo. Un movimiento directo con algún punto fijo es una rotación en torno a una recta ℓ (el conjunto de puntos fijos).

Un movimiento inverso con algún punto fijo es o bien una reflexión sobre un plano Π (el conjunto de puntos fijos) o la composición de una rotación seguida por una reflexión sobre de un Π plano ortogonal al eje de rotación ℓ. En el segundo caso, existe un único punto fijo $O = \ell \cap \Pi$.

Hay tres tipos de movimientos sin puntos fijos. El primer tipo es una traslación no trivial. El segundo tipo es la composición de una rotación seguido por una traslación no trivial paralela al eje de rotación. Tal movimiento rígido es directo, y que se llama un "movimiento helicoidal". El tercer tipo es la composición de una reflexión sobre un plano seguido de una traslación no trivial de vector paralelo a la dirección del plano de reflexión.

Vamos a detallar estas afirmaciones en los párrafos siguientes.

Clasificación de los movimientos en el plano euclídeo

Utilizando la expresión canónica, dada en la Proposición 5.77, de la matriz asociada a una transformación ortogonal sobre una espacio vectorial euclídeo, respecto a una base ortonormal convenientemente elegida, los casos posibles en dimensión dos para la transformación ortogonal asociada a un movimiento en el plano euclídeo son:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \text{ con } a^2+b^2=1, b \neq 0.$$

A esta expresiones se puede llegar con facilidad (para dimensión 2) sin más que considerar $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ la matriz de la aplicación lineal asociada al movimiento e imponerle (al ser ortogonal) que $A\,{}^{\mathrm{t}}A = I$ y ${}^{\mathrm{t}}AA = I$; resultando las ecuaciones:

$$a^2+b^2=1,\ a^2+c^2=1,\ a^2+d^2=1,\ b^2+d^2=1,\quad ac+bd=0,\ ab+cd=0.$$

Si consideramos coordenadas homogéneas en el plano deducidas de la referencia ortonormal tomada, las posibles ecuaciones $\lambda X' = MX$ de un movimiento $((x,y) \mapsto f(x,y))$ son:

$$(\mathbf{1}^{\mathrm{o}}) : \lambda \begin{pmatrix} x'^0 \\ x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ \beta & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x \\ y \end{pmatrix},\ (\mathbf{2}^{\mathrm{o}}) : \lambda \begin{pmatrix} x'^0 \\ x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ \beta & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x \\ y \end{pmatrix},$$

$$(\mathbf{3^o}) : \lambda \begin{pmatrix} x'^0 \\ x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & -1 & 0 \\ \beta & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x \\ y \end{pmatrix}, \quad (\mathbf{4^o}) : \lambda \begin{pmatrix} x'^0 \\ x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & a & b \\ \beta & -b & a \end{pmatrix} \begin{pmatrix} x^0 \\ x \\ y \end{pmatrix}.$$

Estos cuatro tipos de afinidades tiene un punto doble (propio o impropio) y la recta del infinito invariante. Estudiando cada una de ellas, averiguaremos si tiene otros elementos fijos (5) y la distribución de los mismos, esto lo podemos hacer hallando los valores propios de la matriz M (raíces del polinomio característico $|M - \lambda I| = 0$) y los rangos de las matrices $M - \lambda I$ (ver [5, pág. 61] o [6, pág. 191]).

Obsérvese que en todo los casos el polinomio característico $|M - \lambda I| = 0$ tiene la raíz $\lambda = 1$.

(**1º**): $\lambda_1 = 1$ es una raíz triple.

Si $(\alpha, \beta) = (0, 0)$, rango$(M - \lambda_1 I) = 0$: Se trata de la <u>identidad</u>.

Si $(\alpha, \beta) \neq (0, 0)$, rango$(M - \lambda_1 I) = 1$: tiene un recta (la del infinito) de puntos dobles y un punto doble (en el infinito). NO tiene puntos fijos propios. Las rectas (paralelas) que pasan por el punto doble tienen por ecuación cartesiana $\beta x - \alpha y = k$, es decir tiene la dirección del vector (α, β). Se trata de la <u>traslación</u>, cuyo vector de traslación es $\overrightarrow{Xf(X)} = (x + \alpha, y + \beta) - (x, y) = (\alpha, \beta)$.

(**2º**): $\lambda_1 = -1$ una raíz simple, $\lambda_2 = 1$ una raíz doble.

Si $\alpha = 0$, rango$(M - \lambda_2 I) = 1$ (homología): Un punto doble (impropio del eje OY) y una recta ℓ ($y = \beta/2$) invariante. Las rectas (paralelas al eje OY) que pasan por el punto doble son invariantes (en particular, la recta del infinito) y los puntos de la recta invariante son todos dobles. Así, la imagen $f(X)$ de un punto X está en la recta que pasa por X y es perpendicular a ℓ; y como el punto medio de X y $f(X)$ está en ℓ ($X + \frac{1}{2}\overrightarrow{Xf(X)} = (x, y) + \frac{1}{2}(x - x, -y + \beta - y) = (x, \frac{\beta}{2})$), se trata de una <u>simetría</u> ortogonal $f = s_\ell$ de eje la recta ℓ ($y = \beta/2$).

Si $\alpha \neq 0$, rango$(M - \lambda_2 I) = 2$ y, por tanto, existen dos puntos dobles (los del infinito de los ejes OX y OY) y dos rectas invariantes (una la del infinito y $\ell : y = \beta/2$). NO tiene puntos fijos propios. El movimiento es la composición $f = t_{\vec{v}} \circ s_\ell$, donde s_ℓ es la <u>simetría</u> (ortogonal) respecto a la recta ℓ y $t_{\vec{v}}$ es la <u>traslación</u> de vector $\vec{v} = (\alpha, 0)$, paralelo al eje de simetría:
$f(x, y) = (t_{\vec{v}}(s_\ell(x, y)) = s_\ell(x, y) + (\alpha, 0) = (x, -y + \beta) + (\alpha, 0) = (x + \alpha, -y + \beta)$.
Se denomina ***simetría con deslizamiento***.

(**3º**): $\lambda_1 = 1$ una raíz simple, $\lambda_2 = -1$ una doble con rango$(M - \lambda_2 I) = 1$.

A la raíz $\lambda_2 = -1$ le corresponde una recta (la del infinito) de puntos dobles y, a la raíz $\lambda_2 = 1$, un punto doble $P(\alpha/2, \beta/2)$. Las rectas que pasa por el punto doble son invariantes. P es el punto medio de X y $f(X)$, ya que:
$\overrightarrow{PX} + \overrightarrow{Pf(X)} = (x - \alpha/2, y - \beta/2) + (-x + \alpha - \alpha/2, -y + \beta - \beta/2) = (0, 0)$.

(5) El comando de Mathematica `Eigensystem` proporciona los valores propios y los puntos fijos de un movimiento dado su matriz asociada

Sea trata de una simetría central de centro $(\alpha/2, \beta/2)$.

(**4º**): $\lambda_1 = 1$, $\lambda_2 = a + bi$, $\lambda_3 = a - bi$; las raíces simples (una sola real).

A la raíz real le corresponde una recta invariante (la del infinito) y un punto doble:

$$P\left(\frac{(a-1)\alpha - b\beta}{2(a-1)}, \frac{b\alpha + (a-1)\beta}{2(a-1)}\right).$$

Como $a^2 + b^2 = 1$, si ponemos $a = \cos\theta$ y $b = \operatorname{sen}\theta$, el movimiento se escribe:

$$\begin{aligned} x' &= x\cos\theta + y\operatorname{sen}\theta + \alpha \\ y' &= -x\operatorname{sen}\theta + y\cos\theta + \beta. \end{aligned}$$

Representa un giro (o rotación) de amplitud θ y centro P.

Si denotamos las ecuaciones el movimiento, en coordenadas cartesianas, por

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

y sólo consideramos puntos propios, podemos resumir los expuesto en el cuadro siguiente:

Movimientos en el plano euclídeo

$\det A = 1$	$\mathrm{tr}A = 2$	Con puntos fijos	Identidad (**1º**)
		Sin puntos fijos	Traslación (**1º**)
	$\mathrm{tr}A = -2$		Simetría central (**3º**)
	$\lvert\mathrm{tr}A\rvert \neq 2$		Giro (**4º**)
$\det A = -1$	Con puntos fijos		Simetría axial (**2º**)
	Sin puntos fijos		Composición de una simetría axial con una traslación paralela al eje de simetría (**2º**)

Cuando el movimiento es un giro, la matriz de su transformación ortogonal asociada tiene traza $\mathrm{tr}A = 2\cos\theta$. Así, la traza permite determinar el coseno del ángulo de giro, de modo que este pueda ser elegido en el intervalo $[0, \pi]$. Lo único que queda determinar es el signo del ángulo, según la orientación que se haya elegido. Para ello, si la orientación está dada por una base $\{\vec{e}_1, \vec{e}_2\}$, consideramos la base $\{\vec{e}_1, \tilde{f}(\vec{e}_2)\}$, si ambas bases tienen la misma orientación, entonces el ángulo se toma con signo positivo; y si tienen orientación opuesta, se toma con signo negativo.

Clasificación movimientos en el espacio euclídeo tridimensional

Siguiendo el mismo procedimiento que en el plano euclídeo, estudiamos los tipos de movimientos en el espacio euclídeo de dimension 3.

En coordenadas homogéneas deducidas de una referencia ortonormal conveniente elegida, las ecuaciones $\lambda X' = MX$ de un movimiento, $(x,y,z) \mapsto f(x,y,z)$, es de uno de los tipos siguientes:

$$(\mathbf{1^o})\ \lambda \begin{pmatrix} x'^0 \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \alpha & 1 & 0 & 0 \\ \beta & 0 & 1 & 0 \\ \gamma & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x \\ y \\ z \end{pmatrix} \qquad (\mathbf{2^o})\ \lambda \begin{pmatrix} x'^0 \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \alpha & 1 & 0 & 0 \\ \beta & 0 & 1 & 0 \\ \gamma & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x \\ y \\ z \end{pmatrix}$$

$$(\mathbf{3^o})\ \lambda \begin{pmatrix} x'^0 \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \alpha & 1 & 0 & 0 \\ \beta & 0 & -1 & 0 \\ \gamma & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x \\ y \\ z \end{pmatrix} \qquad (\mathbf{4^o})\ \lambda \begin{pmatrix} x'^0 \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \alpha & -1 & 0 & 0 \\ \beta & 0 & -1 & 0 \\ \gamma & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x \\ y \\ z \end{pmatrix}$$

$$(\mathbf{5^o})\ \lambda \begin{pmatrix} x'^0 \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \alpha & 1 & 0 & 0 \\ \beta & 0 & a & b \\ \gamma & 0 & -b & a \end{pmatrix} \begin{pmatrix} x^0 \\ x \\ y \\ z \end{pmatrix} \qquad (\mathbf{6^o})\ \lambda \begin{pmatrix} x'^0 \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \alpha & -1 & 0 & 0 \\ \beta & 0 & a & b \\ \gamma & 0 & -b & a \end{pmatrix} \begin{pmatrix} x^0 \\ x \\ y \\ z \end{pmatrix}$$

Obsérvese que todo los casos el polinomio característico $|M - \lambda I| = 0$ tiene la raíz $\lambda = 1$.

($\mathbf{1^o}$): $\lambda_1 = 1$ es una raíz cuádruple.

Si $(\alpha,\beta,\gamma) = (0,0,0)$, rango$(M - \lambda_1 I) = 0$: Se trata de la identidad.
Si $(\alpha,\beta,\gamma) \neq (0,0,0)$, rango$(M - \lambda_1 I) = 1$: se tiene un plano (el del infinito) de puntos dobles y un punto (en el infinito) doble, $(0:\alpha:\beta:\gamma)$. El vector $\overrightarrow{Xf(X)} = (\alpha,\beta,\gamma)$, con lo que el movimiento es una traslación. NO hay puntos fijos propios.

($\mathbf{2^o}$): $\lambda_1 = -1$ una raíz simple, $\lambda_2 = 1$ una raíz triple.

A la raíz simple le corresponde un punto doble (el del infinito del eje OZ) y un plano doble $(\Pi : z = \gamma/2)$.
Si $(\alpha,\beta) = (0,0)$, rango$(M - \lambda_2 I) = 1$. Rectas y planos paralelos al eje OZ son invariantes y todos los puntos del plano Π son dobles. El punto medio de X y $f(X)$,

$$X + \frac{1}{2}\overrightarrow{Xf(X)} = (x,y,z) + \frac{1}{2}(x+\alpha-x, y+\beta-y, -z+\gamma-z) = (x+\frac{\alpha}{2}, y+\frac{\beta}{2}, \frac{\gamma}{2}) \in \Pi.$$

Se trata de una simetría (ortogonal) respecto al plano Π, s_Π: simetría especular.
Si $(\alpha,\beta) \neq (0,0)$, rango$(M - \lambda_2 I) = 2$: La recta del infinito del plano $(\Pi : z = \gamma/2)$ es invariante; también son invariantes los planos paralelos al eje OZ y al vector $\vec{\boldsymbol{v}} = (\alpha,\beta,0)$. NO hay puntos dobles propios. El movimiento es la composición, $t_{\vec{\boldsymbol{v}}} \circ s_\Pi$, de una simetría especular seguida de una traslación de vector $\vec{\boldsymbol{v}} = (\alpha,\beta,0)$ (paralelo al plano de simetría):

$$f(x,y,z) = (t_{\vec{\boldsymbol{v}}}(s_\Pi(x,y,z)) = s_\Pi(x,y,z) + (\alpha,\beta,0) =$$
$$= (x,y,-z+\gamma) + (\alpha,\beta,0) = (x+\alpha, y+\beta, -z+\gamma).$$

($\mathbf{3^o}$): $\lambda_1 = 1$ una raíz doble, $\lambda_2 = -1$ una raíz doble.

Si $\alpha = 0 \Rightarrow$ rango$(M - \lambda_1 I) = 2$. Los puntos de la recta $\ell : y = \beta/2,\ z = \gamma/2$ son dobles y los planos $\Pi_k : x = k$ (perpendiculares a ℓ) son invariantes.

Si $\alpha = 0 \Rightarrow \text{rango}(M - \lambda_2 I) = 2$. Los puntos de la recta del infinito de los planos paralelos al plano YZ son dobles y los planos, Π_ℓ, que pasan por la recta ℓ $(y = \beta/2, z = \gamma/2)$ son invariantes.

De estos resultados se sigue que las rectas intersección de los planos Π_k y Π_ℓ (que cortan a la recta ℓ) son invariantes; es más, el punto medio de X y $f(X)$ está en ℓ, en efecto:

$$X + \frac{1}{2}\overrightarrow{Xf(X)} = (x, y, z) + \frac{1}{2}(x + \alpha - x, -y + \beta - y, -z + \gamma - z) =$$
$$= (x + \frac{\alpha}{2}, \frac{\beta}{2}, \frac{\gamma}{2}) \in \ell.$$

Se trata de una simetría (ortogonal) respecto a la recta ℓ, s_ℓ: simetría axial.

Si $\alpha \neq 0 \Rightarrow \text{rango}(M - \lambda_1 I) = 3$. El punto del infinito del eje OX es doble y el plano del infinito es invariante.

Si $\alpha \neq 0 \Rightarrow \text{rango}(M - \lambda_2 I) = 2$. La recta del infinito del plano $x = \alpha/2$ (y de todos los paralelos a éste) tiene sus puntos dobles. Los planos que pasan por la recta ℓ $(y = \beta/2, z = \gamma/2)$ son invariantes.

El movimiento (que NO tiene puntos fijos propios) es la composición $t_{\vec{v}} \circ s_\ell$, de una simetría axial seguida de una traslación de vector $\vec{v} = (\alpha, 0, 0)$ (paralelo a la recta de simetría):

$$f(x, y, z) = (t_{\vec{v}}(s_\ell(x, y, z)) = s_\ell(x, y, z) + (\alpha, 0, 0) =$$
$$= (x, -y + \beta, -z + \gamma) + (\alpha, 0, 0) = (x + \alpha, -y + \beta, -z + \gamma).$$

(**4º**): $\lambda_1 = 1$ una raíz simple, $\lambda_2 = -1$ una raíz triple.

A la raíz $\lambda_1 = 1$ le corresponde el punto doble $P(\alpha/2, \beta/2, \gamma/2)$ y el plano del infinito invariante.

Si la raíz es $\lambda_2 = -1$ ocurre que $\text{rango}(M - \lambda_2 I) = 1$, y le corresponde el plano del infinito de puntos dobles y los planos que contiene al punto $P(\alpha/2, \beta/2, \gamma/2)$ son invariantes.

El punto medio de X y $f(X)$ es P:

$$\overrightarrow{PX} + \overrightarrow{Pf(X)} = (x - \frac{\alpha}{2}, y - \frac{\beta}{2}, z - \frac{\gamma}{2} + (-x + \alpha - \frac{\alpha}{2}, -y + \beta - \frac{\beta}{2}, -z + \gamma - \frac{\gamma}{2} = (0, 0, 0).$$

Sea trata de una simetría central de centro $(\alpha/2, \beta/2, \gamma/2)$.

(**5º**): $\lambda_1 = 1$ una raíz doble, $\lambda_2 = a + bi$ y $\lambda_3 = a - bi$ raíces simples.

Si $\alpha = 0 \Rightarrow \text{rango}(M - \lambda_1 I) = 2$. Son dobles los puntos de la recta ℓ (paralela a OX):

$$(a - 1)y + bz + \beta = 0, \quad -by + (a - 1)z + \gamma = 0.$$

Además, los planos Π_k $(x = k)$, paralelos a YZ, son invariantes.

Si trasladamos el sistema rectangular tomado al punto doble $P = \ell \cap \Pi_k$:

$$P\left(k, \frac{(a - 1)\beta - b\gamma}{2(a - 1)}, \frac{b\beta + (a - 1)\gamma}{2(a - 1)}\right),$$

las ecuaciones del movimiento quedan

$$x' = x, \quad y' = ay + bz, \quad z' = -by + az.$$

Con lo que la restricción al plano Π_k es un giro de centro P. Obsérvese que la amplitud de este giro es la misma en cualquier plano Π_k al cual se haga la restricción.

El movimiento en cuestión es, pues, un giro (o rotación), g_ℓ, de eje ℓ.

Si $\alpha \neq 0 \Rightarrow \text{rango}(M - \lambda_1 I) = 3$. A la raíz $\lambda_1 = 1$ le corresponde, ahora, un solo punto doble (punto del infinito del eje coordenado OX) y un único plano invariante (plano del infinito): NO hay puntos dobles finitos.

Se trata de la composición de un giro alrededor de la recta ℓ,

$$(a-1)y + bz + \beta = 0, \quad -by + (a-1)z + \gamma = 0,$$

con la traslación de vector (paralelo al eje de giro) $\vec{\boldsymbol{v}} = (\alpha, 0, 0)$:

$$f(x,y,z) = (t_{\vec{\boldsymbol{v}}}(g_\ell(x,y,z)) = g_\ell(x,y,z) + (\alpha,0,0) =$$
$$= (x, ay+bz+\beta, -by+az+\gamma) + (\alpha,0,0) = (x+\alpha, ay+bz+\beta, -by+az+\gamma).$$

El movimiento recibe el nombre de movimiento helicoidal.

(**6º**): $\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = a+bi$ y $\lambda_4 = a - bi$ raíces simples.

El punto doble correspondiente a la raíz $\lambda_1 = 1$ es $(\alpha/2, 0, 0)$ y el plano invariante, el del infinito.

El punto doble correspondiente a la raíz $\lambda_1 = -1$ es el del infinito del eje coordenado OX y el plano invariante es $\Pi : \; x = \alpha/2$.

En consecuencia, el eje coordenado OX (recta que une los puntos dobles) queda invariante. El movimiento es la composición de un giro alrededor del eje $\ell = OX$, seguido de una simetría respecto al plano Π $(x = \alpha/2)$: $f = s_\Pi \circ g_\ell = g_\ell \circ s_\Pi$,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \alpha & -1 & 0 & 0 \\ \beta & 0 & a & b \\ \gamma & 0 & -b & a \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \alpha & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \beta & 0 & a & b \\ \gamma & 0 & -b & a \end{pmatrix}.$$

A este movimiento se le denomina simetría alabeada o simetría rotacional.

Denotando las ecuaciones el movimiento, en coordenadas cartesianas, por

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix},$$

y sólo consideramos puntos propios, podemos resumir los expuesto en el cuadro siguiente:

Movimientos en el espacio euclídeo tridimensional

$\det A = 1$	$\mathrm{tr}A = 3$	Con puntos fijos	.... Identidad (**1º**)
		Sin puntos fijos	.... Traslación (**1º**)
	$\mathrm{tr}A = -1$	Con puntos fijos	..... Simetría axial (**3º**)
		Sin puntos fijos	. Simetría axial seguida de una traslación (**3º**)
	$\mathrm{tr}A \neq -1, 3$	Con puntos fijos	.... Giro (**5º**)
		Sin puntos fijos	.... Movimiento helicoidal (**5º**)
$\det A = -1$	$\mathrm{tr}A = 1$	Con puntos fijos	Simetría especular (**2º**)
		Sin puntos fijos	.... Simetría especular seguida de una traslación (**2º**)
	$\mathrm{tr}A = -3$		.... Simetría central (**4º**)
	$\mathrm{tr}A \neq -3, 1$		.... Simetría rotacional (**6º**)

Cuando el movimiento es un giro respecto a una recta, la matriz de su transformación ortogonal asociada

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \operatorname{sen}\theta \\ 0 & -\operatorname{sen}\theta & \cos\theta \end{pmatrix},$$

tiene traza $\mathrm{tr}A = 1 + 2\cos\theta$. Así, la traza permite determinar el coseno del ángulo de giro, de modo que este pueda ser elegido en el intervalo $[0, \pi]$. Lo único que queda determinar es el signo del ángulo, según la orientación que se haya elegido. Para ello, si la orientación está dada por una base ortonormal $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, consideramos la base $\{\vec{u}, \vec{v}, \tilde{f}(\vec{v})\}$, con $\vec{u}$ vector propio correspondiente a la raíz 1, $\vec{v}$ vector independiente con $\vec{u}$. Si ambas bases tienen la misma orientación, entonces el ángulo se toma positivo ($\theta = \arccos((\mathrm{tr}A - 1)/2)$); y si tienen orientación opuesta, se toma con signo negativo.

Cuando el movimiento es un giro seguido de una simetría especular (simetría rotacional), ocurre que $\mathrm{tr}A = -1 + 2\cos\theta$ y el signo del ángulo se toma siguiendo el mismo criterio que como se acaba de exponer en el caso del giro.

Determinación de las ecuaciones de un movimiento

En el desarrollo precedente hemos resuelto el problema de la determinación de los elementos geométricos característicos de un movimiento conocida la matriz de éste y de paso clasificarlo, pero queda por estudiar el problema recíproco, acaso más interesante aún, de hallar las matrices que corresponde a un movimiento definido por sus características geométricas.

De este problema vamos a tratar dos casos particulares fundamentales, tanto en el plano como el espacio tridimensional, que nos sirven para determinar fácilmente todos los demás: giro alrededor de un punto y simetría axial en el plano y giro alrededor de una recta y simetría respecto a un plano en el espacio.

Omitimos los cálculos que nos llevan a obtener las ecuaciones de estos movimientos y sólo expondremos sus ecuaciones.

Caso del plano euclídeo

— Ecuaciones del giro de centro en el punto de coordenadas (a, b) y amplitud θ son:

$$\begin{aligned} x' &= (x-a)\cos\theta - (y-b)\operatorname{sen}\theta + a \\ y' &= (x-a)\operatorname{sen}\theta + (y-b)\cos\theta + b. \end{aligned}$$

— Ecuaciones de la simetría axial respecto a la recta dada por su ecuación normal $x\cos\alpha + y\operatorname{sen}\alpha - p = 0$, siendo p la distancia de la recta al origen de coordenadas y α el ángulo que forma la normal a la recta con el eje OX:

$$\begin{aligned} x' &= -x\cos 2\alpha - y\operatorname{sen} 2\alpha + 2p\cos\alpha \\ y' &= -x\operatorname{sen} 2\alpha + y\cos 2\alpha + 2p\operatorname{sen}\alpha. \end{aligned}$$

Si la ecuación de la recta viene de la forma $ax + by + c = 0$, las ecuaciones de la simetría axial son:

$$\begin{aligned} x' &= \frac{1}{a^2+b^2}\left((b^2-a^2)x - 2aby - 2ac\right) = x - \frac{2a(ax+by+c)}{a^2+b^2}, \\ y' &= \frac{1}{a^2+b^2}\left(-2abx + (a^2-b^2)y - 2bc\right) = y - \frac{2b(ax+by+c)}{a^2+b^2}. \end{aligned}$$

Caso del espacio euclídeo tridimensional

— Ecuaciones del giro alrededor de la recta que pasa por el punto $P_0(a,b,c)$ y tiene la dirección del vector unitario de componentes (u,v,w):

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}=\begin{pmatrix} u^2+\left(v^2+w^2\right)\cos t & uv(1-\cos t)-w\operatorname{sen} t & uw(1-\cos t)+v\operatorname{sen} t \\ uv(1-\cos t)+w\operatorname{sen} t & v^2+\left(w^2+u^2\right)\cos t & vw(1-\cos t)-u\operatorname{sen} t \\ uw(1-\cos t)-v\operatorname{sen} t & vw(1-\cos t)+u\operatorname{sen} t & w^2+\left(u^2+v^2\right)\cos t \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
$$-2\operatorname{sen}\frac{t}{2}\begin{pmatrix} (cv-bw)\cos\frac{t}{2}+\left(u(bv+cw)-a\left(v^2+w^2\right)\right)\operatorname{sen}\frac{t}{2} \\ (aw-cu)\cos\frac{t}{2}+\left(v(cw+au)-b\left(w^2+u^2\right)\right)\operatorname{sen}\frac{t}{2} \\ (bu-av)\cos\frac{t}{2}+\left(w(au+bv)-c\left(u^2+v^2\right)\right)\operatorname{sen}\frac{t}{2} \end{pmatrix}.$$

Un caso particular de giros constituyen las simetrías axiales, para las que, siendo $\theta=\pi$, las ecuaciones anteriores se expresan por:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}=\begin{pmatrix} 2u^2-1 & 2uv & 2uw \\ 2uv & 2v^2-1 & 2vw \\ 2uw & 2vw & 2w^2-1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}-2\begin{pmatrix} u(bv+cw)-a\left(v^2+w^2\right) \\ v(cw+au)-b\left(w^2+u^2\right) \\ w(au+bv)-c\left(u^2+v^2\right) \end{pmatrix}.$$

— Ecuaciones de la simetría especular respecto al plano de ecuación $px+qy+rz+s=0$:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}=\begin{pmatrix} \frac{-p^2+q^2+r^2}{p^2+q^2+r^2} & -\frac{2pq}{p^2+q^2+r^2} & -\frac{2pr}{p^2+q^2+r^2} \\ -\frac{2pq}{p^2+q^2+r^2} & \frac{p^2-q^2+r^2}{p^2+q^2+r^2} & -\frac{2qr}{p^2+q^2+r^2} \\ -\frac{2pr}{p^2+q^2+r^2} & -\frac{2qr}{p^2+q^2+r^2} & \frac{p^2+q^2-r^2}{p^2+q^2+r^2} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}-2\begin{pmatrix} \frac{ps}{p^2+q^2+r^2} \\ \frac{qs}{p^2+q^2+r^2} \\ \frac{rs}{p^2+q^2+r^2} \end{pmatrix}$$

Estas ecuaciones, si se toma $p^2+q^2+r^2=1$, se escriben en la forma:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}=\begin{pmatrix} 1-2p^2+q^2+r^2 & -2pq & -2pr \\ -2pq & 1-2q^2 & -2qr \\ -2pr & -2qr & 1-2r^2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}-2\begin{pmatrix} ps \\ qs \\ rs \end{pmatrix}$$

En los siguientes ejemplos vamos a utilizar, como gran ayuda en los cálculos, un paquete de cálculo simbólico.

6.40 Ejemplo.- *Ecuaciones del movimiento helicoidal composición del giro de eje la recta que pasa por el punto $P(2/9,19/2,0)$ y tiene la dirección del vector $\vec{\boldsymbol{v}}=(1,4,1)$ con la traslación $t_{\vec{\boldsymbol{w}}}$ de vector*

$$\vec{\boldsymbol{w}}=(-2/81,-8/81,-2/91).$$

La función personal `matrizGiro3D` de MATHEMATICA permite determinar la matriz de un giro en el espacio tridimensional, dado el eje de giro (punto y dirección) y la amplitud del giro.

```
matrizGiro3D[P_,d_,t_]:=
   Module[{aa,x0,y0,z0,v1,v2,v3,mG,mGH},{x0,y0,z0}=P;
   {v1,v2,v3}=d/Sqrt[d.d]; mG=Array[aa,{3,4},{1,0}];
   aa[1,1]=v1^2(1-Cos[t]])+Cos[t]; aa[2,2]=v2^2(1-Cos[t])+Cos[t];
   aa[3,3]=v3^2(1-Cos[t])+Cos[t]; aa[1,2]=v1*v2(1-Cos[t])-v3*Sin[t];
   aa[2,1]=v1*v2(1-Cos[t])+v3*Sin[t];
   aa[1,3]=v1*v3(1-Cos[t])+v2*Sin[t];
   aa[3,1]=v1*v3(1-Cos[t])-v2*Sin[t];
```

```
    aa[2,3]=v2*v3(1-Cos[t])-v1*Sin[t];
    aa[3,2]=v2*v3(1-Cos[t])+v1*Sin[t];
    {aa[1,0],aa[2,0],aa[3,0]}=
      -(Take[mG,3,-3]-IdentityMatrix[3]).{x0,y0,z0};
    mGH={{1,0,0,0}};
    Do[mGH=Insert[mGH,mG[[i]],i+1],{i,3}];
       mGH];

matrizGiro3D::usage= "matrizGiro3D[P,v,t], determina la matriz del
giro en el espacio tridimensional de amplitud \[Theta], respecto a
la recta que pasa por el punto P y tiene la dirección del vector v
(NO NECESARIAMENTE UNITARIO).";
```

En el Ejemplo que nos ocupa, la instrucción

`M=matrizGiro3D[{2/9,-19/9,0},{1,4,1},π]`

nos que la matriz del giro es:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{110}{81} & -\frac{8}{9} & \frac{4}{9} & \frac{1}{9} \\ -\frac{46}{81} & \frac{4}{9} & \frac{7}{9} & \frac{4}{9} \\ \frac{74}{81} & \frac{1}{9} & \frac{4}{9} & -\frac{8}{9} \end{pmatrix}$$

Para determinar, a partir de esta matriz, los elementos característicos del movimiento, la instrucción `Eigensystem[Take[M,-3,-3]]` , nos da las raíces de polinomio característico de la matriz A de la transformación ortogonal asociada al movimiento, así como los vectores que generan los subespacios directores de la variedades lineales invariantes.

$$\{\{-1,-1,1\}, \quad \{(-1,0,1),(-4,1,0),(1,4,1)\}\}.$$

El resultado es que el polinomio característico $|A-\lambda I| = 0$ tiene la raíz doble -1 y la raíz simple 1.

Este es uno de los casos del tipo (**3º**), en la clasificación de movimientos en el espacio euclídeo tridimensional (pág. 144), cuando tiene puntos fijos: una simetría axial (el giro es de 180°).

A la raíz 1, corresponde el vector de componentes $(1,4,1)$ (el tercero de la lista), que genera el subespacio director de la variedad invariante: una recta, que es el eje de giro.

A la raíz -1, corresponden los vectores de componentes $(-1,0,1)$ y $(-4,1,0)$ (los dos primeros de la lista), que genera el subespacio director de las variedad invariantes: planos, perpendiculares al eje de giro. Se tiene que:

$$(-1,0,1)\times(-4,1,0) = (-1,-4,-1).$$

Para obtener la matriz del movimiento pedido, debemos componer con la traslación de vector $\vec{\boldsymbol{w}} = (-2/81,-8/81,-2/91)$. Para ello le sumamos a la matriz obtenida la siguiente:

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ -\frac{2}{81} & 0 & 0 & 0 \\ -\frac{8}{81} & 0 & 0 & 0 \\ -\frac{2}{81} & 0 & 0 & 0 \end{pmatrix},$$

resultando la matriz:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{4}{3} & -\frac{8}{9} & \frac{4}{9} & \frac{1}{9} \\ -\frac{2}{3} & \frac{4}{9} & \frac{7}{9} & \frac{4}{9} \\ \frac{8}{9} & \frac{1}{9} & \frac{4}{9} & -\frac{8}{9} \end{pmatrix}.$$

6.41 Ejemplo.- *Ecuaciones de la simetría especular respecto al plano de ecuación $2x + y + 2z - 3 = 0$.*

La función del usuario `simetriaEspecular` de MATHEMATICA nos facilita las cuentas:

```
simetriaEspecular[p_,q_,r_,s_]:=Module[{sc}, sc=p^2+q^2+r^2;
    {{1,0,0,0},
     {-2p*s/sc,  (-p^2+q^2+r^2)/sc,-2p*q/sc,-2p*r/sc},
     {-2q*s/sc, -2p*q/sc, (p^2-q^2+r^2)/sc, -2q*r/sc},
     {-2r*s/sc, -2p*r/sc, -2q*r/sc, (p^2+q^2-r^2)/sc}}

simetriaEspecular::usage= "simetriaEspecular[A,B,C,D],
 matriz asociada a la simetría respecto al plano Ax+By+Cz+D=0.";
```

`simetriaEspecular[2,1,2,-3]`, nos da la matriz M asociada a la simetría especular, $X' = AX + B$ con $M = \begin{pmatrix} 1 & 0 \\ B & A \end{pmatrix}$:

$$\begin{pmatrix} 1 \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{4}{3} & \frac{1}{9} & -\frac{4}{9} & -\frac{8}{9} \\ \frac{2}{3} & -\frac{4}{9} & \frac{7}{9} & -\frac{4}{9} \\ \frac{4}{3} & -\frac{8}{9} & -\frac{4}{9} & \frac{1}{9} \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \\ z \end{pmatrix}.$$

Comprobemos que se trata de una simetría especular de acuerdo con la clasificación de movimientos en el espacio euclídeo tridimensional, que hemos dado (pág. 144); así, como cuál es el elemento geométrico que los caracteriza (plano de simetría).

```
M={{1,0,0,0},{4/3,1/9,-4/9,-8/9},{2/3,-4/9,7/9,-4/9},{4/3,-8/9,-4/9,1/9}};
A=Take[M,-3,-3]; (* Extrae las 3 últimas filas y las 3 últimas columnas de M *)
Det[A]
  -1

Tr[A]
  1

Eigensystem[M]  (* Raíces del polinomio característico de M y los
            puntos que generan las variedades correspondientes *)
   {{-1,1,1,1},   {{0,1,1/2,1}, {2/3,0,0,1},{1/3,0,1,0},{2/3,1,0,0}}}

Take[%,-1,-3] (* Puntos que generan la variedad correspondiente a
                la raíz 1 *)
  {{2/3,0,0,1},{1/3,0,1,0},{2/3,1,0,0}}

MatrixRank[%] (* Los tres puntos son independientes: es un plano *)
```

```
  3

Append[%%,{x0,x,y,z}]
  {{2/3,0,0,1}, {1/3,0,1,0}, {2/3,1,0,0}, {x0,x,y,z}}

Factor[Det[%]] (* Primer miembro de la ecuación, en coordenadas homogéneas,
             del plano *)
   1/3 (-2 x + 3 x0 - y - 2 z)
```

Hemos obtenido que el determinate de la matriz, A, de la transformación ortogonal asociada a la afinidad es igual a 1, que su traza es $tr(A) = -1$ y que la variedad lineal invariante correspondiente a la raíz triple $\lambda = 1$ es el plano $-2x - y - 2z + 3 = 0$, es decir, tiene puntos fijos. Todo esto nos lleva, al consultar el cuadro de clasificación, a que es, en efecto, una simetría respecto al plano invariante calculado.

6.42 Ejemplo[Teorema de Hjelmslev].- *Cuando todos los puntos P de una recta están relacionados mediante un movimiento con todos los puntos P' de otra recta, los puntos medios de los segmentos PP' son distintos y alineados o todos ellos coinciden.*

Una verificación de este resultado, puramente algebraica, puede hacerse considerando la clasificación de movimientos en el plano (pág. 142).

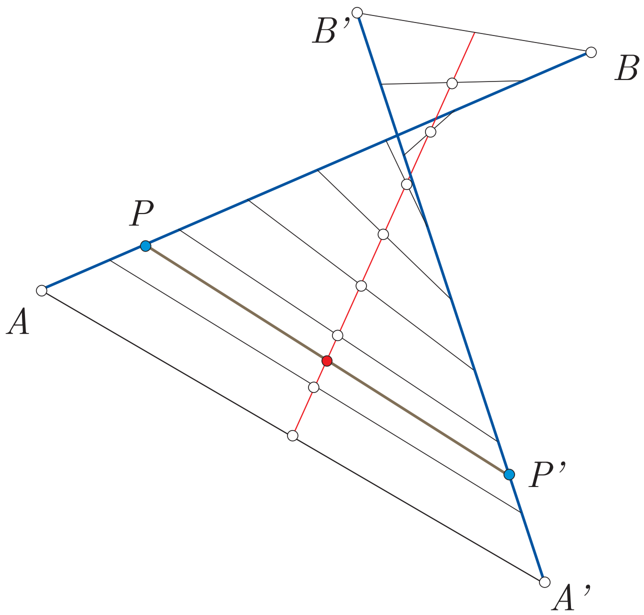
Sea un movimiento en el plano euclídeo real de ecuación matricial $X' = AX + B$ y tomemos una recta determinada por dos puntos arbitrario P y Q. El punto medio de un punto genérico de la recta PQ, $P + t\overrightarrow{PQ} = P + t(Q - P)$, y de su imagen está en una recta de punto genérico:

$$\frac{1}{2}(AP + P + B) + \frac{1}{2}t\big(A(Q - P) + (Q - P)\big).$$

Estos puntos se reducen a un único punto P_0 si y sólo si

$$AP + P + B = 2P_0, \qquad A(Q - P) = -(Q - P).$$

Esto quiere decir que -1 es una raíz del polinomio característico de A y que subespacio asociado a esta raíz es todo $\mathbb{R}^2$ (pues, esto ocurre para cualquier para de puntos del plano). Por consiguiente, la raíz es múltiple. Estamos en el caso (**3^o**) de la clasificación de movimientos en el plano, es decir, se trata de una simetría central; su centro de simetría es el punto $P_0 = \frac{1}{2}B$.



Applet-Cabri http://webpages.ull.es/users/amontes/cabri/g1-20.htm

APÉNDICE A
Planos y rectas en el espacio ordinario

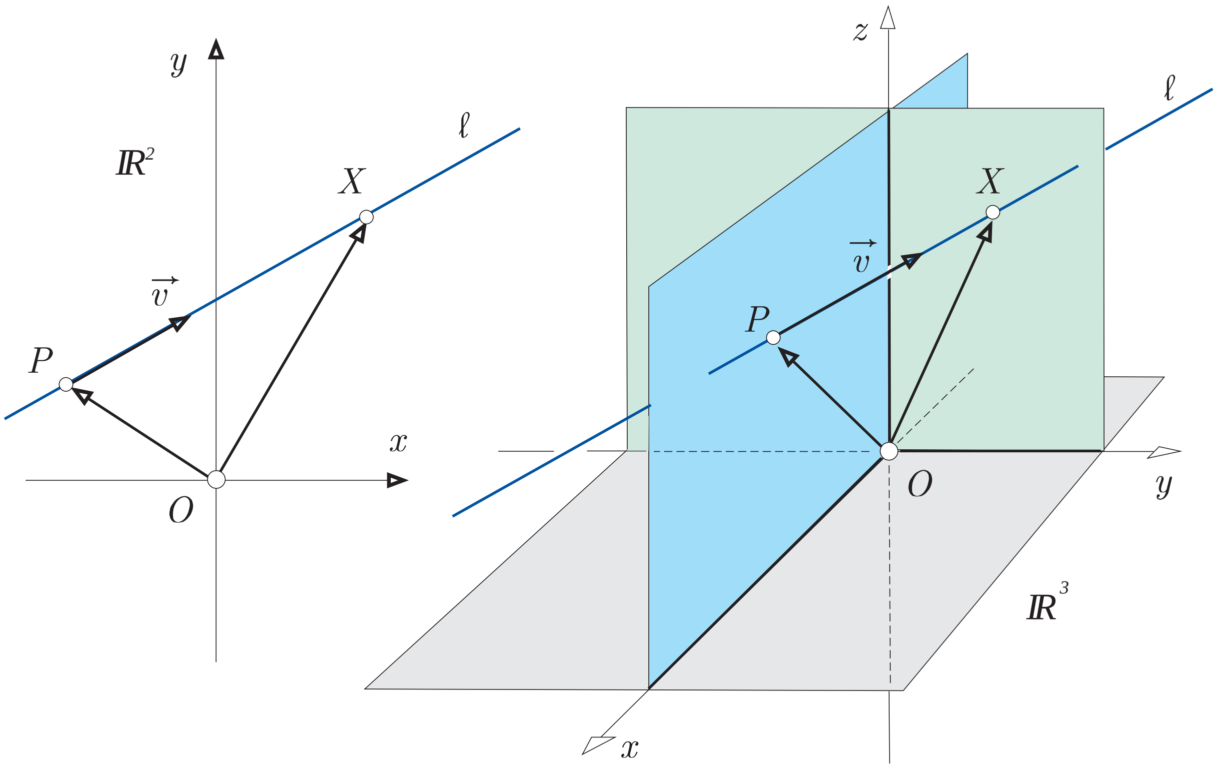
En este Apéndice se hará una exposición, no muy rigurosa desde el punto de vista axiomático, sobre los conceptos de rectas y planos en el espacio ordinario, que nos servirá para resolver ciertos problemas geométricos concretos y como motivación para el desarrollo teórico de los temas anteriores.

A1 Ecuaciones de rectas y planos

Una recta ℓ en el plano o en el espacio ordinarios queda determinada por un punto por donde pasa, P, y por un vector, $\vec{\boldsymbol{v}}$, que determina su dirección (vector director de la recta). Así, un punto genérico de X de ℓ es el extremo del vector desde el origen O del sistema de coordenadas que se haya adoptado:

$$\overrightarrow{OX} = \overrightarrow{OP} + t\vec{\boldsymbol{v}}. \tag{A1.1}$$

Al variar $t \in I\!R$, el punto X recorre la recta ℓ.

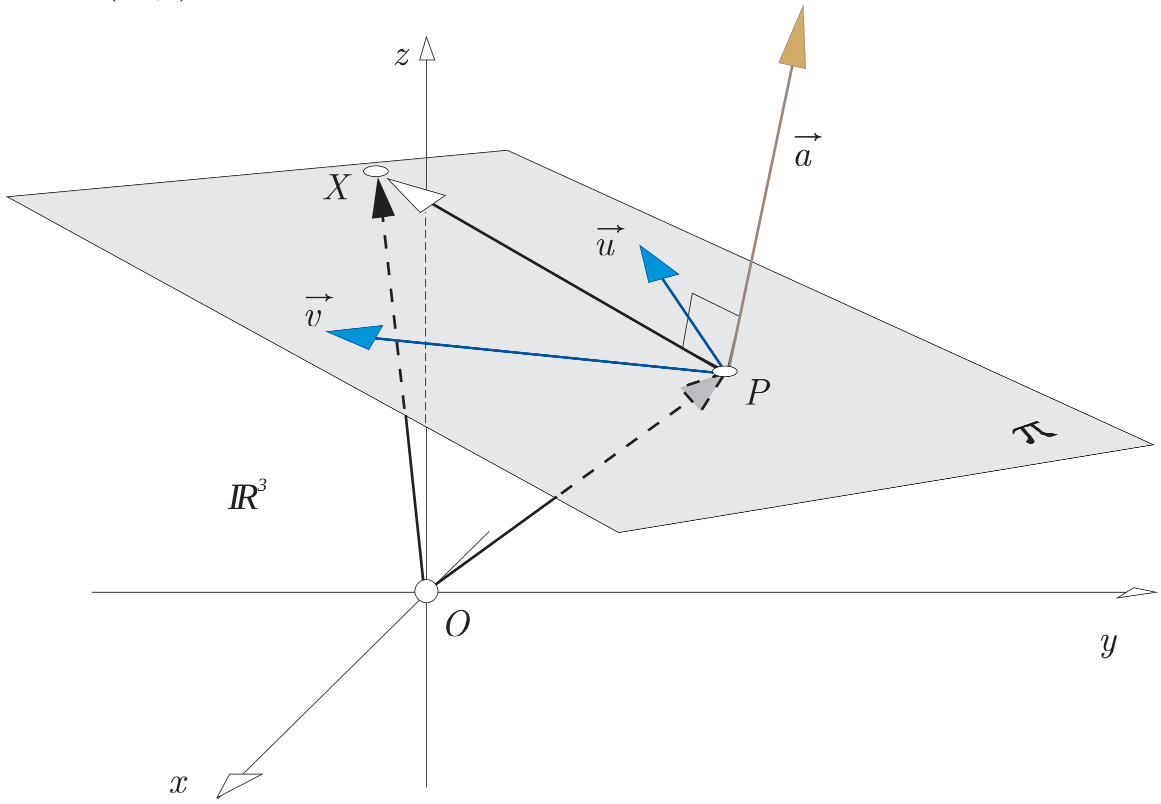


Un plano π en el espacio ordinario queda determinado por un punto P contenido en él y por un vector (director del plano) de dirección perpendicular a

la de toda recta contenida en π, ó por dos vectores independientes de direcciones las de dos rectas contenidas en π. Por lo que, si $\vec{\boldsymbol{u}}$ y $\vec{\boldsymbol{v}}$ son dos vectores aplicados en el punto P y con direcciones las de dos rectas contenidas en π, para cualquier punto X de π se tiene que:

$$\overrightarrow{OX} = \overrightarrow{OP} + \lambda\vec{\boldsymbol{u}} + \mu\vec{\boldsymbol{v}}. \tag{A1.2}$$

Al variar $(\lambda, \mu) \in I\!R^2$, el punto X recorre el plano π.



Si $\vec{\boldsymbol{a}}$ es un vector perpendicular a π, para un punto X de π se debe verificar que:

$$\overrightarrow{PX} \cdot \vec{\boldsymbol{a}} = 0, \tag{A1.3}$$

donde ”$\cdot$” denota el producto escalar de vectores en el espacio vectorial $I\!R^3$.

De las ecuaciones vectoriales obtenidas, que determinan si un punto X está en una recta o en un plano, vamos a deducir otras ecuaciones en términos de las coordenadas de los puntos y las componentes de los vectores, que intervienen en ellas.

• En el caso del plano ordinario, poniendo $P(p, q)$, $X(x, y)$ y $\vec{\boldsymbol{v}} = (v_1, v_2)$, de la ecuación (A1.1), igualando componente a componente, resulta la ecuación paramétrica de la recta:

$$x = p + tv_1, \qquad y = q + tv_2.$$

Eliminando t entre estas dos ecuaciones se llega a la ecuación continua de la recta:

$$\boxed{\frac{x-p}{v_1} = \frac{y-q}{v_2}}$$

Finalmente, operando y pasando todo a primer miembro se obtiene la **ecuación general o implícita** de la recta, en la forma siguiente:

$$Ax + By + C = 0.$$

El par (A, B) son las componentes de un vector perpendicular a la recta.

Si la recta no pasa el origen de coordenadas ($C \neq 0$), de esta ecuación se obtiene la ecuación de la recta en su forma simétrica:

$$\frac{x}{a} + \frac{y}{b} = 1,$$

siendo a y b la abscisa y ordenada en el origen de los puntos de corte de la recta la recta con los ejes.

• En el caso del espacio ordinario, si se parte de la ecuación vectorial del plano (A1.3), poniendo $P(p, q, r)$, $X(x, y, z)$ y $\vec{\boldsymbol{a}} = (a_1, a_2, a_3)$, se llega a:

$$\boxed{a_1(x - p) + a_2(y - q) + a_3(z - r) = 0}$$

Agrupando se obtiene la **ecuación general o implícita** del plano:

$$Ax + By + Cz + D = 0.$$

La terna (A, B, C) son las componentes de un vector perpendicular al plano.

A esta misma ecuación se llega también partiendo de la ecuación (A1.2), poniendo $\vec{\boldsymbol{u}} = (u_1, u_2, u_3)$, $\vec{\boldsymbol{v}} = (v_1, v_2, v_3)$ e igualando componente a componente, resultando la **ecuación paramétrica** del plano:

$$x = p + \lambda u_1 + \mu v_1, \qquad y = q + \lambda u_2 + \mu v_2, \qquad z = r + \lambda u_3 + \mu v_3.$$

Para hallar los valores únicos de λ y μ que corresponden a un punto $X(x, y, z)$ del plano, obtenidos del sistema de ecuaciones anterior de tres ecuaciones y dos incógnitas, ha de ocurrir que el determinate siguiente se anule:

$$\begin{vmatrix} x - p & y - q & z - r \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = 0.$$

Otra vía, para obtener la ecuación del plano, del que se conocen tres puntos $P_i(x_i, y_i, z_i), (i = 1, 2, 3)$, es tener en cuenta que sus coordenadas han de satisfacer a su ecuación general $Ax + By + Cz + D = 0$; por lo que se pueden determinar A, B, C y D, sabiendo que deben ser solución de las ecuaciones:

$$\begin{aligned} Ax \;\; + By \;\; + Cz \;\; + D &= 0 \\ Ax_1 + By_1 + Cz_1 + D &= 0 \\ Ax_2 + By_2 + Cz_2 + D &= 0 \\ Ax_3 + By_3 + Cz_3 + D &= 0 \end{aligned}$$

Y la condición para que exista solución es que:

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0,$$

que desarrollada nos da la ecuación del plano.

Para la recta en el espacio ordinario, que pasa por el punto $P(p, q, r)$ y vector director $\vec{\boldsymbol{v}} = (v_1, v_2, v_3)$, su ecuación paramétrica es:

$$x = p + tv_1, \qquad y = q + tv_2, \qquad z = r + tv_3.$$

Y eliminado el parámetro t entre estas tres ecuaciones, obtenemos las ecuaciones:

$$\boxed{\frac{x-p}{v_1} = \frac{y-q}{v_2} = \frac{z-r}{v_3}}$$

Tomando dos pares de estas ecuaciones y operando, resulta que la ecuación de una recta en el espacio viene determinada por la intersección de dos planos de ecuaciones:

$$Ax + By + Cz + D = 0, \qquad A'x + B'y + C'z + D' = 0.$$

Si los datos que determinan una recta o un plano en el espacio son dados de manera diferente a los expuestos, basta con determinar, a partir de ellos, los elementos que sa han usado aquí para determinar sus ecuaciones. Así, por ejemplo, si de un plano se dan tres puntos no alineados, $P_i(x_i, y_i, z_i), (i = 1, 2, 3)$, se pueden considerar el punto $P = P_1$ y los vectores, contenidos en el plano, $\vec{\boldsymbol{u}} = \overrightarrow{P_1P_2}$ y $\vec{\boldsymbol{v}} = \overrightarrow{P_1P_3}$, y utilizar la ecuación (A1.2) para obtener la ecuación del plano, o bien tomar el vector $\vec{\boldsymbol{a}}$ perpendicular al plano, determinado por el producto vectorial $\vec{\boldsymbol{a}} = \overrightarrow{P_1P_2} \times \overrightarrow{P_1P_2}$, y utilizar la ecuación (A1.3).

A.1. Ejemplo.- *Dadas dos rectas no paralelas y que no se cortan, ℓ_1 y ℓ_2, en el espacio ordinario, determinar la recta ℓ perpendicular a ambas y que las corte.*

Supongamos que las rectas ℓ_1 y ℓ_2 vienen dadas por un punto y un vector director, P_i y $\vec{\boldsymbol{v}}_i$ $(i = 1, 2)$. Un vector director de la recta ℓ es el producto vectorial $\vec{\boldsymbol{a}} = \vec{\boldsymbol{v}}_1 \times \vec{\boldsymbol{v}}_2$.

Los puntos X del plano π_1 que contiene a las rectas ℓ y ℓ_1, verifican:

$$\overrightarrow{P_1X} \cdot (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{v}}_1) = 0.$$

Análogamente, los puntos X del plano π_2 que contiene a las rectas ℓ y ℓ_2, verifican:

$$\overrightarrow{P_2X} \cdot (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{v}}_2) = 0.$$

La recta ℓ pedida es la intersección de los planos π_1 y π_2. Si $P_i(p_i, q_i, r_i)$ y $\vec{\boldsymbol{v}}_i = (u_i, v_i, w_i)$, para $i = 1, 2$, las ecuaciones de los planos son:

$$\begin{aligned} \pi_1: &\quad \mathfrak{S}\left(u_2(v_1^2 + w_1^2) - u_1(v_1v_2 + w_1w_2)\right)(x - p_1) = 0, \\ \pi_2: &\quad \mathfrak{S}\left(u_1(v_2^2 + w_2^2) - u_2(v_1v_2 + w_1w_2)\right)(x - p_2) = 0. \end{aligned}$$

A.2. Ejemplo.- *Determinar la recta perpendicular al eje OZ y a la recta de ecuaciones $x - z - 2 = 0$, $y - z + 1 = 0$, y que corta a ambas.*

Podemos utilizar directamente las fórmulas dadas en el Ejemplo A.1. para determinar dos planos que la determinan, tomando:

$$Q_1(0,0,0),\ \vec{v}_1 = (0,0,1), \quad Q_2(3,0,1),\ \vec{v}_2 = (1,0,-1) \times (0,1,-1) = (1,1,1).$$

$$\begin{array}{ll} \pi_1: & (u_2(v_1^2 + w_1^2) - u_1(v_1v_2 + w_1w_2))(x - p_1) + \cdots = x + y = 0, \\ \pi_2: & (u_1(v_2^2 + w_2^2) - u_2(v_1v_2 + w_1w_2))(x - p_2) + \cdots = -x - y + 2z + 1 = 0. \end{array}$$

O bien, paso a paso, para este caso concreto: La recta perpendicular a ambas rectas dadas, tiene la dirección del vector

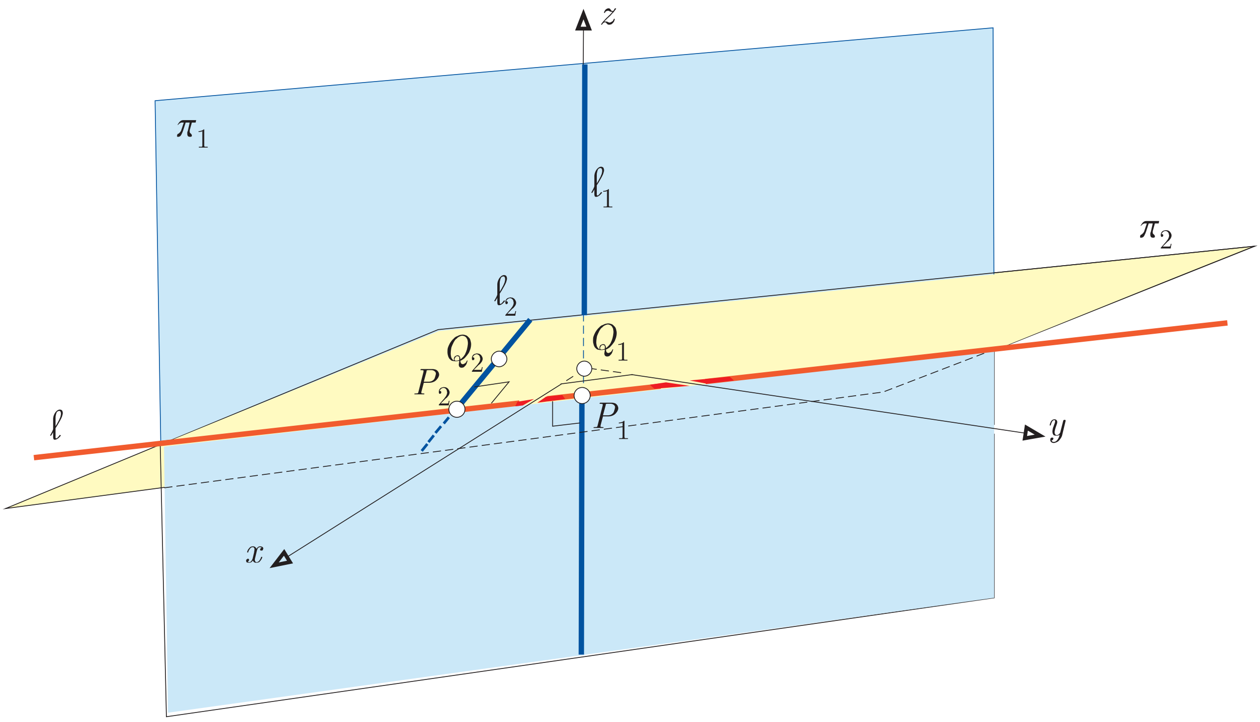
$$\vec{v} = (0,0,1) \times (1,1,1) = (-1,1,0).$$

Un plano, π_1, que contiene a la perpendicular pedida y al eje $\ell_1 = OZ$ es perpendicular al vector $\vec{v}_1 \times \vec{v} = (0,0,1) \times (-1,1,0) = (1,1,0)$ y pasa por $Q_1(0,0,0)$.

Otro plano, π_2, que contiene a la perpendicular pedida y a la segunda recta, ℓ_2 es perpendicular al vector $\vec{v}_2 \times \vec{v} = (1,1,1) \times (-1,1,0) = (1,1,-2)$ y pasa por $Q_2(3,0,1)$.

Luego,

$$\begin{array}{ll} \pi_1: & (x,y,z) \cdot (1,1,0) = x + y = 0, \\ \pi_2: & (x-3, y, z-1) \cdot (1,1,-2) = x + y - 2z - 1 = 0. \end{array}$$



http://webpages.ull.es/users/amontes/jview/g1-01.html

Los puntos $P_i = \ell \cap \ell_i (i = 1,2)$, pies de la perpendicular común a las rectas ℓ_1 y ℓ_2, son $P_1(0,0,-1/2)$ y $P_2(3/2,-3/2,-1/2)$. Por lo que la distancia entre las dos rectas es $\|\overrightarrow{P_1P_2}\| = 3/\sqrt{2}$.

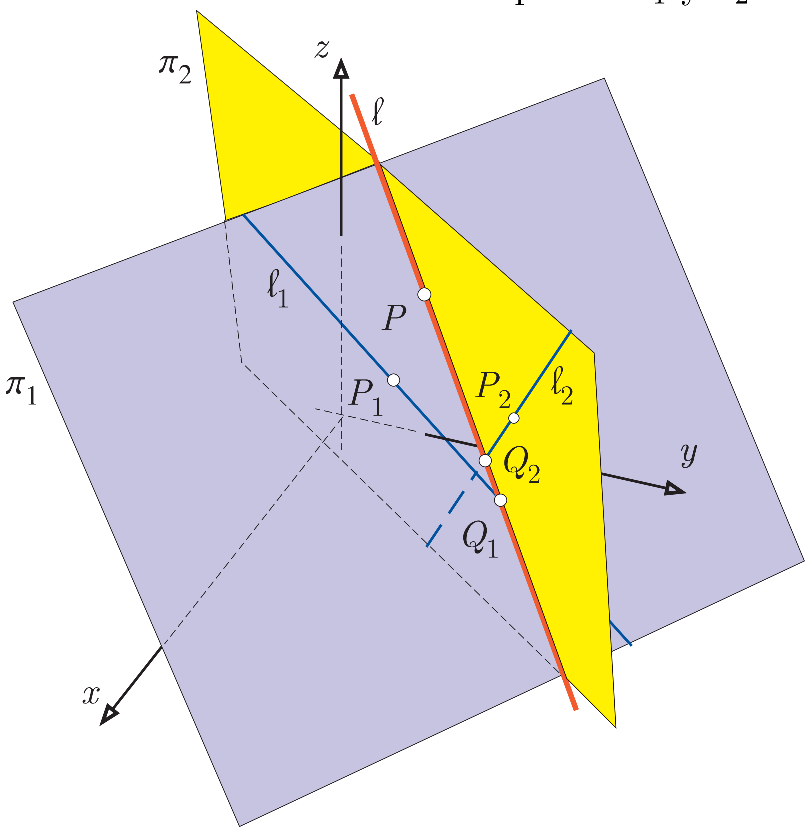
A.3. Ejemplo.- *Recta que pasa por $P(1,1,2)$ y se apoya en la recta ℓ_1 : $x+y+z=2, 2x-y+z=3$ y en la recta ℓ_2 que pasa por $P_2(3,0,1)$ y que tiene la dirección del vector $(1,-1,2)$.*

La recta ℓ_1 pasa por el punto $P_1(1,0,1)$ y tiene vector director $(1,1,1)\times(2,-1,1)=(2,1,-3)$.

El plano que pasa por $P(1,1,2)$ y contiene a la recta ℓ_1 es perpendicular al vector $(2,1,-3)\times\overrightarrow{P_1P}=(2,1,-3)\times(0,-1,-1)=(-4,2,-2)$, por lo que su ecuación es $\pi_1 :\ 2x-y+z-3=0$.

El plano que pasa por $P(1,1,2)$ y contiene a la recta ℓ_2 es perpendicular al vector $(1,-1,2)\times\overrightarrow{P_2P}=(1,-1,2)\times(-2,1,1)=(-3,-5,-1)$, por lo que su ecuación es $\pi_2 :\ 3x+5y+z-10=0$.

La recta pedida ℓ es la intersección de los planos π_1 y π_2.



http://webpages.ull.es/users/amontes/jview/g1-02.html

Los puntos $Q_i=\ell\cap\ell_i (i=1,2)$ son $Q_1(5/2,3/4,-5/4)$ y $Q_2(11/5,4/5,-3/5)$.

A.4. Ejemplo.- *Determinar el circuncentro del triángulo cuyos vértices son los puntos $A(2,-1,0)$, $B(1,1,-3)$ y $C(0,1,2)$.*

Los planos perpendiculares a cada lado en su punto medio se cortan en una recta que pasa por el circuncentro del triángulo $\widehat{ABC}$, que es el punto que equidista de los tres vértices.

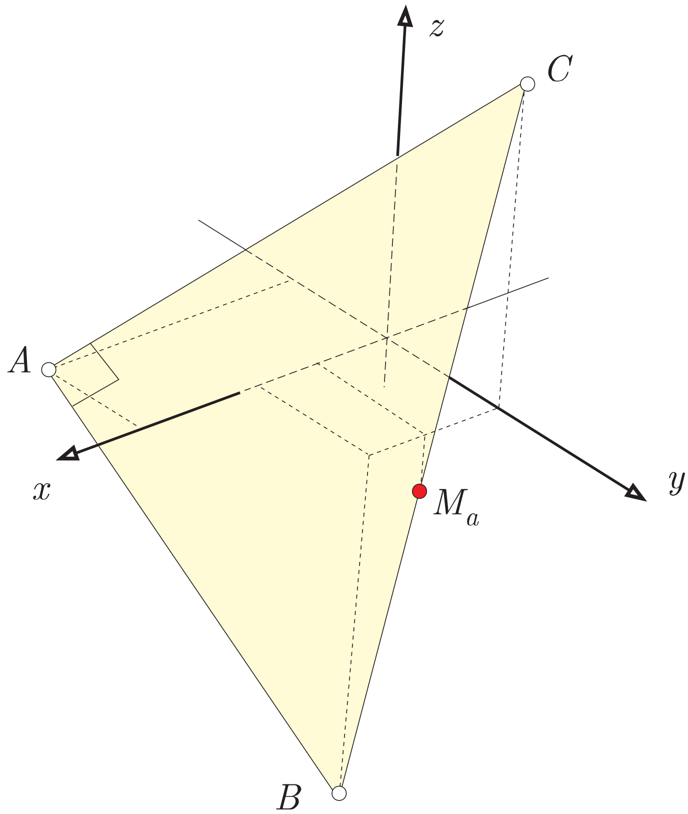
Los puntos medios de los lados de los lados BC, CA y AB son, respectivamente, $M_a(1/2,1,-1/2), M_b(1,0,1)$ y $M_c(3/2,0,-3/2)$.

Las ecuaciones de los planos mediatrices a los lados BC, CA y AB son, respectivamente,

$$\pi_a :\ -x+5z+3=0, \quad \pi_b :\ x-y-z=0, \quad \pi_c :\ -x+2y-3z-3=0.$$

Por ejemplo, el primero de ellos se obtiene sabiendo que pasa por M_a y es perpendicular al vector $\overrightarrow{BC} = (-1, 0, 5)$.

El plano que contiene al triángulo $\widehat{ABC}$ es $\pi : \ 5x + 4y + z - 6 = 0$. Luego, la intersección de los plano $\pi_a \cap \pi_b \cap \pi$ es el punto $\left(\frac{1}{1}, 1, -\frac{1}{2}\right)$, circuncentro de $\widehat{ABC}$.



http://webpages.ull.es/users/amontes/jview/g1-05.html

El producto escalar de los vectores $\overrightarrow{AB} = (-1, 2, -3)$ y $\overrightarrow{AC} = (-2, 2, 2)$ es cero; luego, el ángulo en el vértice A es recto; por tanto, el circuncentro es el punto medio M_a del lado BC.

Bibliografía

[1] **M. Anzola; J. Caruncho**.- *Problemas de Álgebra, Tomo 6. Geometría Afín y Euclídea*. 1981

[2] **Juan de Burgos**.- *Curso de álgebra y geometría*. Alhambra. Madrid, 1988

[3] **Manuel Castellet, Irene Llerena**.- *Algebra Lineal y Geometría*. Ed. Reverté. Barcelona, 2000

[4] **Jean Gallier**.- *Geometric Methods and Applications For Computer Science and Engineering*, (Texts in Applied Mathematics). Springer-Verlag, TAM Vol. 38 (2000)

[5] **Angel Montesdeoca**- *Geometría proyectiva. Cónicas y Cuádricas.* Colección Textos Universitarios. Gobierno de Canarias. 2001.

Disponible en: http://webpages.ull.es/users/amontes/apuntes/gdh.pdf

[6] **Luis Angel Santaló**.- *Geometría proyectiva*. Eudeba. Buenos Aires. 1966

[7] **Sebastià Xambó Descamps**.- *Geometría*. Ed. UPC. Barcelona, 1997

Bibliografía adicional

[8] **Michele Audin**.- *Geometry.* Springer Universitext, 2002.

[9] **Marcel Berger**.- *Géométrie 1* Universitext, Springer Verlag, 1990

[10] **Marcel Berger**.- *Géométrie 2* Universitext, Springer Verlag, 1990

[11] **H.S.M. Coxeter**.- *Introduction to Geometry*, Wiley, 1989

[12] **A. Doneddu**.- *Curso de Matemáticas. Complementos de Geometría Algebraica.* Colección Ciencia y Técnica. Aguilar. 1980.

[13] **Jean Fresnel**.-*Methodes Modernes en Geometrie* Hermann, 1996

[14] **Roger Godement**.- *Álgebra* Ediciones Tecnos, 1971 Springer Verlag, 1990

[15] **Melvin Hausner**.- *A vector space approach to geometry*, Dover, 1998

[16] **D. Hilbert; S. Cohn-Vossen**.- *Geometry And The Immagination* AMS Chelsea, 1932

[17] **Jacqueline Lelong-Ferrand; Jean-Marie Arnaudiès**.- *Curso de matemáticas. Tomo III. Geometría y Cinemática.* Editorial Reverté S.A. Barcelona 1982.

[18] **Dan Pedoe**.- *Geometry. A comprehensive course*, Dover, 1988

[19] **Pedro Puig Adam**.- *Curso de Geometría Métrica* (2 vols.). Biblioteca Matemática S.L. Madrid 1973.

[20] **Ernst Snapper; Robert J. Troyer**.-*Metric Affine Geometry*, Dover, 1989

SÍMBOLOS

Índice alfabético